PA 1017787

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

June 02, 2003

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE UNDER 35 USC 111.

APPLICATION NUMBER: ***60/386,818***
FILING DATE: ***June 06, 2002***

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)

By Authority of the
COMMISSIONER OF PATENTS AND TRADEMARKS

L. Edelen

L. EDELEN
Certifying Officer

06-10-02 60[redacted]818.0606

Please type a plus sign(+) inside this box → [+]

PTO/SB/16 (02-01)
Approved for use through 10/31/2002 OMB 0651-0032
U S. Patent and Trademark Office; U S DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PROVISIONAL APPLICATION FOR PATENT COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53 (c).

Express Mail Label No. ET916844137US

INVENTOR(S)		
Given Name (first and middle [if any])	Family Name or Surname	Residence (City and either State or Foreign Country)
Arturas	Petronis	Toronto, Ontario, CANADA

☐ *Additional inventors are being named on the separately numbered sheets attached hereto*

TITLE OF THE INVENTION (280 characters max)

DETECTION OF EPIGENTIC ABNORMALITIES AND DIAGNOSTIC METHOD BASED THEREON

Direct all correspondence to: **CORRESPONDENCE ADDRESS**

☐ Customer Number [ ] *Type Customer Number here* →

Place Customer Number *Bar Code Label here*

OR

X	Firm or Individual Name	ARMSTRONG, WESTERMAN & HATTORI, LLP			
Address		502 Washington Avenue			
Address		Suite 220			
City	Towson	State	MD	ZIP	21204
Country	U.S.A.	Telephone	410-337-2295	Fax	410-337-2296

ENCLOSED APPLICATION PARTS *(check all that apply)*

- X Specification *Number of pages* 35
- X Drawing(s) *Number of sheets* 66
- ☐ Application Data Sheet. See 37 CFR 1.76
- ☐ CD(s), Number
- ☐ Other (specify)

METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT

- ☐ Applicant claims small entity status. See 37 CFR 1.27.
- X A Check or money order is enclosed to cover the filing fees — FILING FEE AMOUNT ($) $160.00
- X The Commissioner is hereby authorized to charge filing fees or credit any overpayment to Deposit Account Number: 02-2839
- ☐ Payment by credit card. Form PTO-2038 is attached.

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.

- X No
- ☐ Yes, the name of the U S Government agency and the Government contract number are. _____

Respectfully submitted,

SIGNATURE Robert M Gamson Date 06/06/02

TYPED or PRINTED NAME Robert M. Gamson REGISTRATION NO. *(if appropriate)* 32,986

TELEPHONE 410-337-2295 Docket Number: 02-153-PPA

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

This collection of information is required by 37 CFR 1.51. The information is used by the public to file (and by the PTO to process) a provisional application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 8 hours to complete, including gathering, preparing, and submitting the complete provisional application to the PTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U S. Patent and Trademark Office, U S. Department of Commerce, Washington, D.C., 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Box Provisional Application, Assistant Commissioner for Patents, Washington, D. C. 20231

ARMSTRONG, WESTERMAN & HATTORI, LLP

BALTIMORE OFFICE
SUITE 220
502 WASHINGTON AVENUE
TOWSON, MARYLAND 21204

PHONE: (410) 337-2295
FACSIMILE: (410) 337-2296
TOLL FREE: 1-877-7EDISON
E-MAIL: lbpatent@mindspring.com

June 6, 2002

VIA EXPRESS MAIL
ET916844137US

To the Honorable Commissioner
of Patents and Trademarks
P. O. Box 2327
Arlington, VA 22202

CERTIFICATE OF TRANSMITTAL
I hereby certify that this correspondence is being deposited with the U.S. Postal Service as Express Mail in an envelope addressed to: Commissioner of Patents and Trademarks, P.O. Box 2327, Arlington, VA 22202.
Date: June 6, 2002
Express Mail Label No. ET916844137US
By: Carolyn A. Bates

Re: Our Docket No. 02-153-PPA

Dear Sir:

Enclosed please find the following:

1. New U.S.A. provisional patent application entitled "DETECTION OF EPIGENTIC ABNORMALITIES AND DIAGNOSTIC METHOD BASED THEREON", including specification, claims and abstract (35 pages), drawings (66 pages) Petronis, Inventor.

2. Form PTO/SB/16 duly executed.

3. Our check No. 00152, in the amount of $160.00, to cover the application filing fee.

4. Our post card. (Please date stamp and return.)

Please address all correspondence to:

ARMSTRONG, WESTERMAN & HATTORI, LLP
502 Washington Avenue
Suite 220
Towson, MD 21204.

The inventor is:

Arturas Petronis
250 College Street
Toronto, Ontario
CANADA M5T 1R8
Citizenship: Canadian

If there are any additional fees required, please charge our Deposit Account No. 02-2839.

Thank you for your cooperation and assistance.

Respectfully submitted,

Robert M. Gamson

RMG/chb
Enclosures

C:\Corel\Office7\WPWin7\Carolyn\pto\02-153ptoltr-ppa.wpd

DETECTION of EPIGENTIC ABNORMALITIES and DIAGNOSTIC METHOD BASED THEREON

The present invention relates to identification of epigenetic abnormalities. More particularly, the present invention relates to diagnosis of diseases based on DNA methylation differences, and identification and isolation of genes that cause such diseases.

BACKGROUND OF THE INVENTION

Substantial progress has been made in recent years with respect to the diagnosis and treatment of diseases in which a single defective gene is responsible. Traditional linkage studies have effectively isolated the causal gene and allowed for the further development of diagnostic tests and furthered research into treatments such as gene therapy for conditions such as cystic fibrosis, Duchennes muscular dystrophy, Huntington's disease and fragile X syndrome. However, similar progress has not been made in diseases caused by mutations in multiple genes. Traditional linkage studies in complex diseases such as schizophrenia, bipolar disorder, cancers and diabetes have only succeeded in isolating chromosome regions, often containing 200-300 genes. The ability to screen such a large number of genes is clearly a time-consuming and daunting task.

Epigenetic mechanisms can be an important factor in complex, multi-factorial diseases such as cancers. Epigenetics refers to modifications in gene expression that are brought about by heritable, but potentially reversible changes in DNA methylation and chromatin structure (Henikoff S, Matzke MA Exploring and explaining epigenetic effects. Trends Genet 1997,13(8):293-5; Siegfried Z, Eden S, Mendelsohn M, Feng X, Tsuberi BZ, Cedar H. DNA methylation represses transcription in vivo. Nat Genet 1999, 22(2):203-206; Gonzalgo, M.L. and Jones, P.A. (1997) Mutagenic and epigenetic effects of DNA methylation. Mutat. Res. 386(2), 107-18; Razin, A. and Shemer, R. (1999) Epigenetic control of gene expression. Results Probl. Cell. Differ. 25, 189-204; Lyko, F.

and Paro, R. (1999) Chromosomal elements conferring epigenetic inheritance. Bioessays 21(10), 824-32). DNA methylation of the binding sites for transcription factors changes the affinity of such factors for regulatory sequences, which affects the transcriptional activity of a gene (Ehrlich M and Ehrlich K (1993) Effect of DNA methylation and the binding of vertebrate and plant proteins to DNA. In: Jost JP and Saluz P (eds) DNA Methylation: Molecular Biology and Biological Significance pp. 145-168. Birkhauser Verlag, Basel, Switzerland; Riggs A, Xiong Z, Wang L, and LeBon JM (1998) Methylation dynamics, epigenetic fidelity and X chromosome structure. In: Wolffe AP (ed) Epigenetics, pp. 214-227. John Wiley & Sons, Chistester). In addition to positional effects of methylated cytosines, density in a gene regulatory region also contributes to gene activity. This type of regulation is mediated by methylated cytosine binding proteins and acetylation of histones (Jones PL, Veenstra GJ, Wade PA, Vermaak D, Kass SU, Landsberger N, Strouboulis J, and Wolffe AP (1998) Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nature Genetics 19: 187-91; Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN, and Bird A (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature 393: 386-9; Robertson KD and Wolffe AP (2000) DNA methylation in health and disease. Nature Review Genet 1:11-9).

Methylation can occur within cytosine-guanosine islands (CpG islands) that are typically between 0.2 to about 1 kb in length and are located upstream of many housekeeping and tissue-specific genes, but may also extend into protein coding regions. Methylation of cytosine residues contained within CpG islands of certain genes has been inversely correlated with gene activity. This could lead to decreased gene expression by a variety of mechanisms including, for example, disruption of local chromatin structure, inhibition of transcription factor-DNA binding, or by recruitment of proteins which interact specifically with methylated sequences indirectly preventing transcription factor binding. Some studies have demonstrated an inverse correlation between methylation of CpG islands and gene expression. Tissue-specific genes are usually unmethylated within the receptive target organ cells but are methylated in the germline and in non-expressing

adult tissues. CpG islands of constitutively-expressed housekeeping genes are normally unmethylated in the germline and in somatic tissues.

US5871917 discloses methods for detecting epigenetic abnormalities comprising: restriction of genomic DNA with a methylation-sensitive restriction enzyme (a restriction enzyme that cleaves an unmethylated site, but does not cleave the same site if it is methylated) that leaves an overhang; ligation of adaptors to the overhangs; PCR amplification with primers directed to the adaptors; followed by a subtractive hybridization to eliminate house keeping genes; and a second round of PCR amplification with a second set of primers directed to a second set of adaptors. A problem with this design is that the method is limited to a restriction enzyme that leaves overhangs and, further, the method is complicated due to the ligation of two sets of adaptors.

WO99/01580 discloses methods for detection of genomic imprinting disorders based on digestion of genomic DNA with methylation-sensitive restriction enzymes and PCR amplification using primers. One embodiment, directed to the detection of unmethylated sequences, requires the use of a restriction enzyme that leaves overhangs and the use of exogenous adaptors, and therefore suffers from similar disadvantages as those described above in regards to US5871917. Another embodiment, directed to the detection of methylated sequences, uses primers directed to endogenous elements such that exogenous adaptors are not required, but these primers are required to be positioned on either side of a methylation-sensitive restriction site. Since a methylation sensitive restriction enzyme will cut an unmethylated site, this method can only be used to amplify the methylated sequences, and cannot produce an unmethylated sequence which will be cut in between the two primers.

It is an object of the present invention to overcome disadvantages of the prior art.

The above object is met by a combination of the features of the main claims. The sub claims disclose further advantageous embodiments of the invention.

SUMMARY OF THE INVENTION

The present invention relates to detection of epigenetic abnormalities and diagnosis of non-Mendelian diseases associated with epigenetic abnormalities, and identification and isolation of genes that cause such diseases.

According to the present invention there is provided a method of detecting an epigenetic abnormality associated with a non-Mendelian disease, the method comprising:

a) extraction of genomic DNA from a sample that exhibits characteristics of a non-Mendelian disease;

b) digestion of the genomic DNA with a methylation-sensitive restriction enzyme to produce a pool of restricted DNA fragments;

c) fractionation of the pool of restricted DNA fragments to obtain DNA fragments of a desired size;

d) amplification of at least a segment of the DNA fragments of a desired size with primers that anneal to an endogenous DNA element to produce a PCR product;

e) cloning of the PCR product into a sequencing vector;

f) sequence determination of the PCR product to obtain a sequence of the PCR product;

g) comparing the sequence against a genomic database to assign a locus for the epigenetic abnormality associated with a non-Mendelian disease.

A non-Mendelian disease is any multi-factorial disease such as schizophrenia, bipolar disorder, cancer, and diabetes.

The sample from which DNA is extracted may be any cell, tissue, organ or other suitable specimen that exhibits characteristics of a non-Mendelian disease. For example, without wishing to be limiting, in an individual suffering from schizophrenia or bipolar disorder a sample may be obtained from brain tissue.

Any endogenous DNA element that is found to have epigenetic abnormalities associated with a non-Mendelian disease can be PCR amplified according to the present invention. In a further aspect, the endogenous DNA element is a multi-copy DNA element. In a still further aspect, the multi-copy DNA element is selected from the group consisting of LINE, SINE, L1, and Alu.

In another aspect, the present invention provides a method of identifying a gene having an epigenetically altered expression pattern that contributes to a non-Mendelian disease in an organism, the method comprising:

a) extraction of genomic DNA from a sample that exhibits characteristics of a non-Mendelian disease;

b) digestion of the genomic DNA with a methylation-sensitive restriction enzyme to produce a pool of restricted DNA fragments;

c) fractionation of the pool of restricted DNA fragments to obtain DNA fragments of a desired size;

d) amplification of at least a segment of the DNA fragments of a desired size with primers that anneal to an endogenous DNA element to produce a PCR product;

e) cloning of the PCR product into a sequencing vector;

f) sequence determination of the PCR product to obtain a sequence of the PCR product;

g) comparing the sequence against a genomic database to assign a locus for said epigenetic abnormality associated with a non-Mendelian disease;

h) searching said database to identify a gene located proximal to said locus;

h) comparing expression patterns of said gene located proximal to said locus within a test sample that exhibits characteristics of said non-Mendelian disease with expression patterns of a corresponding gene within a control sample to identify said gene having an epigenetically altered expression pattern.

Genes can be identified in accordance with the present invention from any eukaryotic organism including, plants and animals, where epigenetic abnormality is associated with the occurrence of non-Mendelian disease.

In yet another aspect, the present invention provides a method of isolating a probe for detecting an epigenetic abnormality associated with a non-Mendelian disease in an animal, said method comprising:

a) extraction of genomic DNA from a sample that exhibits characteristics of a non-Mendelian disease;

b) digestion of said genomic DNA with a methylation-sensitive restriction enzyme to produce a pool of restricted DNA fragments;

c) fractionation of said pool of restricted DNA fragments to obtain DNA fragments of a desired size;

d) amplification of at least a segment of said DNA fragments of a desired size with primers that anneal to an endogenous DNA element to produce a PCR product;

f) using said PCR product as said probe to detect said epigenetic abnormality associated with a non-Mendelian disease in another sample.

This summary does not necessarily describe all necessary features of the invention but that the invention may also reside in a sub-combination of the described features.

BRIEF DESCRIPTION OF THE DRAWINGS

These and other features of the invention will become more apparent from the following description in which reference is made to the appended drawings wherein:

FIGURE 1 shows the localization of the cloned Alu elements.

FIGURE 2 shows coding genes that are located in the vicinity (within 100,000 bp) of cloned Alu elements.

FIGURE 3 shows sequences of the cloned Alu elements.

DESCRIPTION OF PREFERRED EMBODIMENT

The invention relates to methods and compositions for identification of epigenetic abnormalities. More particularly, the present invention relates to diagnosis of diseases based on DNA methylation differences and identification of genes that cause such diseases. The present invention provides methods and compositions for detecting and isolating DNA sequences which are abnormally or differentially methylated in a diseased cell type when compared to a normal cell type.

Traditional linkage studies in complex diseases such as schizophrenia, bipolar disorder, cancers and diabetes have only succeeded in isolating chromosome regions, often containing 200-300 genes. The ability to screen such a large number of genes is clearly a time-consuming and daunting task. The present invention provides a short-cut in determining which genes within a 200-300 gene region are in fact responsible for the onset of a major disease such as diabetes, schizophrenia, cancers, or bipolar disorder. According to the present invention differentially modified, endogenous multi-copy DNA elements can act as markers for genes which are dys-regulated. Epigenetic analysis of so called "junk" DNA leads to a 'short-cut' in identification of specific genes, dys-regulation of which increases the risk to major disease.

The following description is of a preferred embodiment by way of example only and without limitation to the combination of features necessary for carrying the invention into effect.

The methylation patterns of DNA from tumor cells are generally different than those of normal cells (Laird et al., DNA Methylation and Cancer, 3 HUMAN MOLECULAR GENETICS 1487, 1488 (1994)). Tumor cell DNA is generally undermethylated relative to normal cell DNA, but selected regions of the tumor cell genome may be more highly methylated than the same regions of a normal cell's genome. Hence, detection of altered methylation patterns in the DNA of a

tissue sample is an indication that the tissue is cancerous. For example, the gene for Insulin-Like Growth Factor 2 (IGF2) is hypomethylated in a number of cancerous tissues, such as Wilm's Tumors, rhabdomyosarcoma, lung cancer and hepatoblastomas (Rainner et al. 362 NATURE 747-49 (1993); Ogawa, et al., 362 NATURE 749-51 (1993); S. Zhan et al., 94 J. CLIN. INVEST. 445-48 (1994); P. V. Pedone et al., 3 HUM. MOL. GENET. 1117-21 (1994); H. Suzuki et al., 7 NATURE GENET 432-38 (1994); S. Rainier et al., 55 CANCER RES. 1836-38 (1995)).

Alteration of methylation may be a key, and common event, in the development of neoplasia and may play at least two roles in tumorigenesis:
1) DNA hypomethylation may cause an increase in proto-oncogene expression or DNA hypermethylation may decrease expression of a tumor supressor which contributes to neoplastic growth; and
2) DNA hypomethylation may change chromatin structure, and induce abnormalities in chromosome pairing and disjunction. Such structural abnormalities may result in genomic lesions, such as chromosome deletions, amplifications, inversions, mutations, and translocations, all of which are found in human genetic diseases and cancer.

While the present invention can be used for detecting any alteration in methylation, the present invention is particularly useful for detecting and isolating DNA fragments that are normally methylated but which, for some reason, are non-methylated in a proportion of cells. Such DNA fragments may normally be methylated for a number of reasons. For example, such DNA fragments may be normally methylated because they contain, or are associated with, genes that are rarely expressed, genes that are expressed only during early development, genes that are expressed in only certain cell-types, and the like.

As used herein, hypomethylation means that at least one cytosine in a CG or CNG di- or tri-nucleotide site in genomic DNA of a given cell-type does not contain CH_3 at the fifth position of the cytosine base. Cell types which may have hypomethylated CGs or CCGs include any cell type which may be expressing a non-housekeeping function. This

includes both normal cells that express tissue-specific or cell-type specific genetic functions, as well as tumorous, cancerous, and similar cell types. Cancerous cell types and conditions which can be analyzed, diagnosed or used to obtaining probes by the present methods include, but are not limited to, Wilm's cancer, breast cancer, ovarian cancer, colon cancer, kidney cell cancer, liver cell cancer, lung cancer, leukemia, rhabdomyosarcoma, sarcoma, and hepatoblastoma.

A method of the present invention is directed to detection of epigenetic abnormalities associated with a non-Mendelian disease and comprises extraction of genomic DNA from a non-Mendelian disease sample, such as diseased tissue or diseased population of cells; hydrolysis of this DNA with methylation-sensitive restriction enzymes, and subsequent fractionation of DNA fragments and purification of DNA fragments of a desired size, for example, but not limited to, shorter than 10 kB. These purified DNA fragments are further subjected to PCR amplification using primers that hybridize to endogenous multi-copy DNA elements including, but not limited to, ALU or L1 elements. After that, PCR products of such elements are cloned and sequenced using standard molecular biology techniques known to the skilled artisan and the resultant sequences are mapped on the genome using any commercially or publicly available human genome database. These cloned multi-copy elements indicate a loci of putative epigenetic abnormality or epigenetic dys-regulation and indicates genes that predispose a patient to a complex, non-Mendelian, multi-factorial disease, such as, but not limited to, cancers, diabetes, schizophrenia, or bipolar disorder.

By the term "non-Mendelian disease" is meant any disease which etiologically requires more than a single genetic abnormality. As such a non-Mendelian disease requires more than one factor, or in other words, is multi-factorial, and may comprise epigenetic alterations or abnormalities.

Epigenetics relates to higher order gene control mechanisms in eukaryotes that activate or repress parts of the genome via changes in chromatin structure. These higher order gene control mechanisms form an important molecular basis of cell differentiation. Any

changes in an organism brought about by alterations in the action of genes, where the changes do not require occurrence of any mutations, are called epigenetic changes. An epigenetic abnormality occurs when an epigenetic change contributes or predisposes normal cells into becoming diseased cells. DNA methylation is an example of an epigenetic mechanism. The term DNA methylation refers to the addition of a methyl group to the cyclic carbon 5 of a cytosine nucleotide. A family of conserved DNA methyltransferases catalyzes this reaction. Normally, DNA methylation can be used, for example, but is not limited to, to methylate the transcription unit of a gene so that the gene is turned off or silenced, and a corresponding protein product is not produced in a particular cell. For instance, one of the two X chromosomes in female mammals is inactivated or silenced by methylation.

DNA is extracted from a non-Mendelian disease sample using standard techniques, known in the art, for isolating DNA from various samples such as cells , tissues, or organs, or other suitable specimens. Standard techniques for isolating DNA have are disclosed in reference textbooks or manuals such as Sambrook, Fritsch, and Maniatis, Molecular Cloning: A Laboratory Manual (1989), Cold Spring Harbor.

DNA may be extracted from any sample that may have epigenetic abnormalities associated with a non-Mendelian disease or any sample that exhibits characteristics of a non-Mendelian disease, for example, but not limited to cells of the following tissues: Epithelial Tissues, Exocrine Glands, Endocrine Glands, Connective Tissues, Adipose Tissue, Cartilage, Bone, Blood, Muscle Tissues comprising Smooth, Skeletal or Cardiac Muscle Tissue, or Nervous Tissue comprising Brain Tissue.

Any methylation-sensitive restriction enzyme may be used for the purposes of this invention. The terms "restriction endonucleases" and "restriction enzymes" refer to bacterial enzymes, each of which cut double-stranded DNA at or near a specific nucleotide sequence. The process of cutting or cleaving the DNA is referred to as restriction digestion. The products of a restriction digestion are referred to as restriction

products. A restriction enzyme used in the present invention may yield restriction products having blunt-ends or overhanging "sticky" ends. Specifically, a restriction enzyme can symmetrically cut both strands of a double stranded DNA fragment to produce a blunt-ended fragment, or a restriction enzyme may assymetrically cleave the two strands of a DNA fragment to produce a DNA fragment that has a single stranded overhang. In general, a methylation-sensitive restriction enzyme used in the present invention will recognize and cleave a non-methylated sequence, while it will not cleave a corresponding methylated sequence. Methylation of plant and mammalian DNA occurs at CG or CNG sequences. This methylation may interfere with the cleavage by some restriction endonucleases. Endonucleases that are sensitive and not sensitive to m^5CG or m^5CNG methylation, as well as isoschizomers of methylation-sensitive restriction endonucleases that recognize identical sequences but differ in their sensitivity to methylation, can be extremely useful for studying the level and distribution of methylation in eukaryotic DNA. Examples of methylation-sensitive restriction enzymes, and corresponding restriction site sequences, that can be used according to the present invention include, but are not limited to: AatII (GACGTC); Bsh1236I (CGCG); Bsh1285I (CGRYCG); BshTI (ACCGGT); Bsp68I (TCGCGA); Bsp119I (TTCGAA); Bsp143II (RGCGCY); Bsu15I (ATCGAT); Cfr10I (RCCGGY); Cfr42I (CCGCGG); CpoI (CGGWCCG); Eco47III (AGCGCT); Eco52I (CGGCCG); Eco72I (CACGTG); Eco105I (TACGTA); EheI (GGCGCC); Esp3I (CGTCTC); FspAI (RTGCGCAY); Hin1I (GRCGYC); Hin6I (GCGC); HpaII (CCGG); Kpn2I (TCCGGA); MluI (ACGCGT); NotI (GCGGCCGC); NsbI (TGCGCA); PauI (GCGCGC); PdiI (GCCGGC); Pfl23II (CGTACG); Psp1406I (AACGTT); PvuI (CGATCG); SalI (GTCGAC); SmaI (CCCGGG); SmuI (CCCGC); TaiI (ACGT); or TauI (GCSGC).

Size fractionation and purification of restricted DNA fragments can be performed by any method known in the art, for example, but not limited to, separation of DNA fragments of a desired size such as fragments of less than 10 kB by centrifugation of a DNA fragment pool through a membrane or other suitable matrix having size exclusion or inclusion properties. Alternatively, a pool of restricted DNA fragments may be separated using agarose of polyacrylamide gel electrophoresis and DNA fragments of a desired size

may be purified using any suitable gel-extraction composition such as glass milk or Quaternary ammonium ions. The desired size limit of the fractionated and isolated DNA fragments depends on the size of the endogenous DNA element that serves as a template for PCR amplification. As such the "DNA fragments of a desired size" can be any size as long as they are larger than, and can therefore comprise the endogenous DNA element.

As used, the terms "amplification," "amplify," or "amplifying," are defined as the production of additional copies of a nucleic acid sequence and is generally carried out using polymerase chain reaction (PCR) or other technologies well known in the art (e.g., Dieffenbach and Dveksler, PCR Primer, a Laboratory Manual, Cold Spring Harbor Press, Plainview NY [1995]). Nucleic acid amplification techniques allow for increasing the concentration of a target or template sequence, or a portion or segment thereof from a mixture of genomic DNA without cloning or purification. A review of current nucleic acid amplification technology can be found in Kwoh et al., 8 Am. Biotechnol. Lab. 14 (1990). In vitro nucleic acid amplification techniques include polymerase chain reaction (PCR), transcription-based amplification system (TAS), self-sustained sequence replication system (3SR), ligation amplification reaction (LAR), ligase-based amplification system (LAS), Q.beta. RNA replication system and run-off transcription. All present and future nucleic acid amplification technology can be incorporated into the present invention.

PCR is a preferred method for DNA amplification. PCR synthesis of DNA fragments occurs by repeated cycles of heat denaturation of DNA fragments, primer annealing onto endogenous sequence elements or exogenous adaptor ends of a DNA fragment or other suitable DNA template, and primer extension. These cycles can be performed manually or, preferably, automatically. Thermal cyclers such as the Perkin-Elmer Cetus cycler are specifically designed for automating the PCR process, and are preferred. The number of cycles per round of synthesis can be varied from 2 to more than 50, and is readily determined by considering the source and amount of the nucleic acid template, the desired yield and the procedure for detection of the synthesized DNA fragment.

PCR techniques and many variations of PCR are known. Basic PCR techniques are described by Saiki et al. (1988 Science 239:487-491) and by K.B. Mullis in U.S. Pat. Nos. 4,683,195, 4,683,202 and 4,800,159, which are incorporated herein by reference.

The conditions generally required for PCR include temperature, salt, cation, pH and related conditions needed for efficient amplification of at least a segment or portion of a DNA fragment template. PCR conditions include repeated cycles of heat denaturation, and incubation at a temperature permitting primer hybridization to an endogenous sequence elements or exogenously ligated adaptors, and copying of the DNA fragment by the amplification enzyme. Heat stable amplification enzymes like the pwo, Thermus aquaticus or Thermococcus litoralis DNA polymerases are commercially available which eliminate the need to add enzyme after each denaturation cycle. The salt, cation, pH and related factors needed for enzymatic amplification activity are available from commercial manufacturers of amplification enzymes.

As provided herein an amplification enzyme is any enzyme which can be used for in vitro nucleic acid amplification, e.g. by the above-described procedures. Amplification enzymes may be thermostable or thermolabile. Such amplification enzymes include pwo, Escherichia coli DNA polymerase I, Klenow fragment of E. coli DNA polymerase I, T4 DNA polymerase, T7 DNA polymerase, Thermus aquaticus (Taq) DNA polymerase, Thermococcus litoralis DNA polymerase, SP6 RNA polymerase, T7 RNA polymerase, T3 RNA polymerase, T4 polynucleotide kinase, Avian Myeloblastosis Virus reverse transcriptase, Moloney Murine Leukemia Virus reverse transcriptase, T4 DNA ligase, E. coli DNA ligase, Vent polymerases, or Q.beta. replicase. Preferred amplification enzymes are the pwo and Taq polymerases. The pwo enzyme is especially preferred because of its fidelity in replicating DNA.

With PCR, it is possible to amplify a single copy of a specific target sequence in genomic DNA to a level detectable by several different methodologies (e.g., hybridization with a labeled probe; incorporation of biotinylated primers followed by avidin-enzyme conjugate detection; incorporation of 32P-labeled deoxynucleotide triphosphates, such

as dCTP or dATP, into the amplified segment). In addition to genomic DNA, any oligonucleotide sequence can be amplified with the appropriate set of primer molecules. In particular, the amplified segments created by the PCR process itself are, themselves, efficient templates for subsequent PCR amplifications.

By the term "primer" is meant an oligonucleotide, whether occurring naturally as in a purified restriction digest or produced synthetically, capable of acting as a point of initiation of synthesis when placed under suitable conditions in which synthesis of a primer extension product that is complementary to a nucleic acid strand is induced. Such suitable conditions comprise nucleotides and an amplification enzyme such as DNA polymerase and a suitable temperature, salt concentration, and pH). The primer is preferably single stranded for maximum efficiency in amplification, but may alternatively be double stranded. If double stranded, the primer is first treated to separate its strands before being used to prepare extension products. The primer must be sufficiently long to prime the synthesis of extension products in the presence of the inducing agent. The exact lengths of the primers will depend on many factors, including temperature, salt concentration , pH, source of primer and the use of the method. The primers of the present invention can hybridize or anneal to a sequence element that is endogenous to a DNA fragment template or the primers can anneal to exogenous adaptor sequence elements that have been ligated to the ends of a DNA fragment template. Preferably, the primers anneal to an endogenous multi-copy DNA sequence element, for example, long or short interspersed nucleotide elements (LINEs or SINEs)..

Endogenous multi-copy DNA sequence elements are repetitive DNA sequences that together are estimated to comprise 30% of total genomic sequences. Present at between 10 - 105 copies per genome these multi-copy elements can be found throughout the euchromatin and have been categorized as:

a) microsatellites / minisatellites (VNTR, DNA 'fingerprints)

b) dispersed-repetitive DNA, mainly transposable elements (LINES/ SINES)

Also includes 'redundant' genes for histones, and ribosomal RNA and proteins, (gene-products present in cell in large numbers).

Many multi-copy DNA elements may be involved in regulation of gene expression as they have been shown to be interspersed within single-copy sequences and have been shown to be located adjacent structural genes.

Long and short interspersed nucleotide elements (LINEs and SINEs), are represented in humans mainly by L1 (Furano AV. The biological properties and evolutionary dynamics of mammalian LINE-1 retrotransposons. Prog Nucleic Acid Res Mol Biol. 2000;64:255-94) and Alu elements (Watson et al., Molecular Biology of the Gene, fourth edition (1987) pp. 669-670), respectively. Both types of elements are considered to be retrotransposable (ie. can replicate via an RNA copy reinserted as DNA by reverse transcription) and they have significant roles in genomic function. The inserted elements can be full length or truncated, or may be rearranged relative to full-length elements.

The most common and best characterised LINE is L1, having the following properties:

- Repeated approximately 50000 times in the human genome (0 5% of total)
- Only about 3000 of these are full length; the remainder are truncated, mostly at the 5' end.
- Full length element is about 6kb in size and contains two open reading frames, one of which encodes a reverse transcriptase.
- AT-rich region is located near the 3' end of the element,
- Element is flanked by two short direct repeats.

The main type of SINE is the Alu family, characterized as follows:

- usually contain a target for the restriction enzyme Alu I;
- 5×10^5 - 10^6 copies in the haploid genome, with an average of one repeat every 4 to 5 kb (1 - 10 % total);
- Often present in the transcription unit of a gene, within introns and occasionally in non-translated regions of the mRNA;

- Generally contain 300bp consensus sequence which consist of two tandem repeats of a 130bp sequence, one of which has a 32bp deletion, as such Alu family members are recognizably related in sequence, but not precisely conserved;
- Elements are flanked by direct repeats;
- Each repeat unit has an AT-rich region that suggests a poly A tail;
- 5' end resembles a pol III promoter region.

LINEs and SINEs both have a poly(A) tail which may act as a template for reverse transcription from nicks made at the site of insertion in the host DNA by a LINE-encoded endonuclease.

Primers of the present invention may be designed according to any L1 or Alu sequence. For example, various analyses (Claverie,J.M. and Makalowski,W. Alu alert, Nature 371, 752 (1994)) indicate that Alu repeats fall into 8 subfamilies, and therefore, 8 ALU consensus sequences have been constituted and added to GenBank as accession numbers U14567, U14568, U14569, U14570, U14571, U14572, U14573 and U14574. A primer of the present invention may be designed in accordance with any of these consensus sequences. For example, the deposited consensus sequence of a subfamily of Alu repeats designated U14570 is as follows:

GGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGA
GGCGGGTGGATCATGAGGTCAGGAGATCGAGACCATCCTGGCTAACAAGG
TGAAACCCCGTCTCTACTAAAAATACAAAAAATTAGCCGGGCGCGGTG

Products of amplification reactions can be subjected to sequence determinations. Amplification products, preferably PCR products, can optionally be cloned into a vector before sequencing. When not cloning a PCR product, an adaptor DNA elements can be ligated to the ends of PCR products, and the PCR products can be sequenced using a primer that anneals to the adaptor element. Cloning, ligation, and sequencing can be performed using standard techniques , such as protocols described in textbooks or manuals such as Sambrook, Fritsch and Maniatis, Molecular Cloning: A Laboratory

Manual, 1989. Also, commercially available kits may be utilized. Another alternative for sequence determination are automated DNA sequencing systems and methods.

Nucleic acid sequences of amplification products isolated according to methods of present are disclosed in Figure 3. The region of the chromosome to which a given sequence is located may be determined by hybridization, including, but not limited to PCR amplification methods, or by database searching.

Hybridization methods and conditions are well known in the art. Nucleic acids that are identical to the provided nucleic acid sequences, bind to the provided nucleic acid sequences (disclosed in Figure 3) under stringent hybridization conditions. By using probes, particularly labeled probes of DNA sequences, one can determine a region of chromosome where a given sequence is located and thereby establish chromosomal loci for epigenetic abnormalities associated with a non-Mendelian disease.

Preferably, hybridization is performed using at least 15 contiguous nucleotides from any sequence identified by the methods of the present invention including, but not limited to, sequences disclosed in Figure 3. The probe will preferentially hybridize with a nucleic acid comprising a complementary sequence to the probe, allowing the identification of the chromosomal region of the nucleic acids of the biological material that uniquely hybridize to the selected probe. Probes of more than 15 nucleotides can be used, e.g. probes of from about 18 nucleotides up to the entire length of the provided nucleic acid sequences, but 15 nucleotides generally represents sufficient sequence for unique identification.

As mentioned above once the sequence (or a portion of the sequence) of a multi-copy DNA element has been isolated, this sequence can be used to map the location of the multi-copy DNA element on a chromosome. Accordingly, nucleic acids of the invention described herein or fragments thereof, can be used to map the location of multi-copy DNA elements of the invention on a chromosome. The mapping of the sequences of

nucleic acids of the invention to chromosomes is an important first step in correlating these sequences with genes associated with disease.

Briefly, sequences of the invention, for example, sequences disclosed in Figure 3, can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp in length) from the sequences of nucleic acids of the invention. These primers can then be used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human sequence corresponding to the sequences of nucleic acids of the invention will yield an amplified fragment.

Somatic cell hybrids are prepared by fusing somatic cells from different mammals (e.g., human and mouse cells). As hybrids of human and mouse cells grow and divide, they gradually lose human chromosomes in random order, but retain the mouse chromosomes. By using media in which mouse cells cannot grow (because they lack a particular enzyme), but in which human cells can, the one human chromosome that contains the gene encoding a needed enzyme, depending on the media, will be retained. By using various media, panels of hybrid cell lines can be established. Each cell line in a panel contains either a single human chromosome or a small number of human chromosomes, and a full set of mouse chromosomes, allowing easy mapping of individual sequences to specific human chromosomes. (D'Eustachio et al. (1983) Science 220:919-924). Somatic cell hybrids containing only fragments of human chromosomes can also be produced by using human chromosomes with translocations and deletions.

PCR mapping of somatic cell hybrids is a rapid procedure for assigning a particular sequence to a particular chromosome. Three or more sequences can be assigned per day using a single thermal cycler. Using the sequences of nucleic acids of the invention to design oligonucleotide primers, sublocalization can be achieved with panels of fragments from specific chromosomes. Other mapping strategies which can similarly be used to map a sequence of a nucleic acid of the invention to its chromosome include in situ hybridization (described in Fan et al. (1990) Proc. Natl. Acad. Sci. USA 87:6223-27), pre-screening with labeled flow-sorted chromosomes,

pre-selection by hybridization to chromosome specific cDNA libraries, and searching of genomic databases.

Once the sequence (or a portion of the sequence) of a multi-copy DNA element has been isolated, this sequence can be used to map the location of the gene on a chromosome by searching a genomic database, for example, but not limited to, a human genome database (www.genome.ucsc.edu/). Several genome databases are also available from Celera Corp. or the National Center for Biotechnology Information (NCBI). Genome databases can be searched by comparing the known query sequence or reference sequence with genomic sequences stored and annotated in a database, and selecting sequences from the database that have a high similarity, preferably greater than 80% similarity, with the query or reference sequence. Sequence similarity is calculated based on a reference sequence, which may be a subset of a larger sequence, such as a conserved motif, coding region, flanking region, etc. A reference sequence will usually be at least about 18 contiguous nucleotides long, more usually at least about 30 nucleotides long, and may extend to the complete sequence that is being compared. Algorithms for sequence analysis are known in the art, such as BLAST, described in Altschul et al., J. Mol. Biol. (1990) 215:403-10.

To determine whether a nucleic acid exhibits similarity with the sequences presented herein, oligonucleotide alignment algorithms may be used, for example, but not limited to a BLAST (GenBank URL: www.ncbi.nlm.nih.gov/cgi-bin/BLAST/, using default parameters: Program: blastn; Database: nr; Expect 10; filter: default; Alignment: pairwise; Query genetic Codes: Standard(1)), BLAST2 (EMBL URL: http://www.embl-heidelberg.de/Services/ index.html using default parameters: Matrix BLOSUM62; Filter: default, echofilter: on, Expect:10, cutoff: default; Strand: both; Descriptions: 50, Alignments: 50), or FASTA, search, using default parameters.

Fluorescence in situ hybridization (FISH) of a DNA sequence to a metaphase chromosomal spread can further be used to provide a precise chromosomal location in one step. Chromosome spreads can be made using cells whose division has been blocked

in metaphase by a chemical, e.g., colcemid that disrupts the mitotic spindle. The chromosomes can be treated briefly with trypsin, and then stained with Giemsa. A pattern of light and dark bands develops on each chromosome, so that the chromosomes can be identified individually. The FISH technique can be used with a DNA sequence as short as 500 or 600 bases. However, clones larger than 1,000 bases have a higher likelihood of binding to a unique chromosomal location with sufficient signal intensity for simple detection. Preferably 1,000 bases, and more preferably 2,000 bases will suffice to get good results at a reasonable amount of time. For a review of this technique, see Verma et al., (Human Chromosomes: A Manual of Basic Techniques (Pergamon Press, New York, 1988)). Sequences of isolated multi-copy DNA elements of the present invention that are shorter than 500 bases can be extended by any suitable technique, for example, a known sequence can be extended by a technique of genomic sequencing using a primer designed according to the known sequence.

Reagents for chromosome mapping can be used individually to mark a single chromosome or a single site on that chromosome, or panels of reagents can be used for marking multiple sites and/or multiple chromosomes. Reagents corresponding to noncoding regions of the genes actually are preferred for mapping purposes. Coding sequences are more likely to be conserved within gene families, thus increasing the chance of cross hybridizations during chromosomal mapping.

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. (Such data are found, for example, in V. McKusick, Mendelian Inheritance in Man, available on-line through Johns Hopkins University Welch Medical Library). The relationship between genes and disease, mapped to the same chromosomal region, can then be identified through linkage analysis (co-inheritance of physically adjacent genes), described in, e.g., Egeland et al. (1987) Nature 325: 783-787.

Probes specific to the nucleic acids of the invention can be generated using a whole or portion of the nucleic acid sequences disclosed in Figure 3. The probes can be

synthesized chemically or can be generated from longer nucleic acids using restriction enzymes. The probes can be labeled, for example, with a radioactive, biotinylated, or fluorescent tag. Preferably, probes are designed based upon an identifying sequence of a nucleic acid of one of Figure 3. More preferably, probes are designed based on a contiguous sequence of one of the subject nucleic acids that remain unmasked following application of a masking program for masking low complexity (e.g., XBLAST) to the sequence., i.e. one would select an unmasked region, as indicated by the nucleic acids outside the poly-n stretches of the masked sequence produced by the masking program. Probes are not only useful for determining chromosomal location of a sequence, but also can be used to determine whether an epigenetic abnormality exists in another sample, for example a test sample obtained from a eukaryotic organism that exhibits symptoms of a non-Mendelian disease.

Once a chromosomal locus has been assigned to a multi-copy DNA element obtained by the present invention, a genomic database or genetic map data can be used to identify one or more genes that are proximal to the assigned chromosomal locus, preferably the identified one or more genes are physically adjacent to the assigned locus. Expression patterns of the genes in a non-Mendelian disease sample can then be compared against the expression pattern of corresponding genes in a control sample to identify a gene having an epigenetically altered expression pattern. The non-Mendelian disease sample and the control sample can be obtained from within the same organism, for example, without wishing to be limiting, expression of a gene within cancerous kidney cells could be compared against expression of a corresponding gene in a non-cancerous kidney cell of the same organism. Alternately, the non-Mendelian disease sample and the control sample can be obtained from different organisms. For example, without wishing to be limiting, expression of a gene in a prefrontal cortex sample from a schizophrenic individual can be compared against expression of a corresponding gene in a prefrontal cortex sample from a different non-schizophrenic individual.

Techniques for determining expression patterns of genes are well known in the art. For example, gene expression patterns can be established using Northern analysis, reporter

constructs such as GFP, quantitative PCR amplification, or DNA chip analysis. If, for example, gene expression within a sample is determined using DNA chips, the mRNA from the sample is extracted, reverse transcribed to the corresponding cDNA, amplified, fluorescently labeled and allowed to hybridize with the sequences on a chip. Sequence-specific labels are captured on the surface of the chip. By reading the fluorescence, one can determine which of the genes were expressed and at what levels. DNA chip analysis is provided by several companies, for example, but not limited to, Affymetrix and Nanogen. DNA chip technology is an effective method for determining expression patterns of genes and semiconductor fabrication technology has allowed for the packing of thousands of gene sequences into square centimeter surfaces. Use of reporter constructs, Northern analysis, and quantitative PCR amplification are equally effective alternatives.

Potential therapeutic approaches.

Detection of epigenetic abnormalities associated with non-Mendelian diseases including, but not limited to schizophrenia, diabetes, cancers and bipolar disorder may lead to innovative DNA modification-based therapies. Recently a compound protein consisting of a DNA methylation enzyme and a zinc-finger protein was constructed (Xu G-L, Bestor TH. Nature Genetics 17: 376-379, 1997). The mechanism of action of the protein consists of the recognition of a specific DNA sequence by the zinc-finger protein that is specific for that sequence and subsequent modification of the surrounding cytosines by DNA modification enzymes. A specific protein with DNA modification enzyme restoring the normal pattern of DNA methylation can be generated. The blood-brain barrier has been a major obstacle for the bloodborne genetic constructs to reach the brain, but a recent study demonstrated that pegylated neutral liposomes, unlike cationic ones, are stable in blood, do not get entrapped in the lung, and are able to efficiently deliver plasmid DNA through the blood brain barrier to the various sections of brain tissue .

The present invention provides methods and compositions for detecting DNA elements that act as a marker for the specific dysfunctional genes and at the same time identify the

specific genes involved in a non-Mendelian diseases. Such information would lead quickly to the development of a diagnostic test for such diseases, that could be incorporated into a diagnostic kit. Further research on specific genes may also lead to treatment options for people suffering from a non-Mendelian disease through either gene therapy work or through targeted drug development.

The heuristic value of epigenetics in non-Mendelian diseases, including schizophrenia, derives from numerous important characteristics of epigenetic regulation of genes (Petronis A. Human morbid genetics revisited: relevance of epigenetics. Trends Genet. 2001 Mar;17(3):142-6). The epigenetic research program indicates that regulation of gene activity is critically important for normal functioning of the genome. Genes, even the ones that carry no mutations or disease predisposing polymorphisms, may be useless or even harmful if not expressed in the appropriate amount, at the right time of the cell cycle, or in the right compartment of the nucleus. Epigenetic mechanisms, more so than DNA sequence-based ones, can explain a series of phenomenological features of a non-Mendelian disease, for example, in the case of, major psychosis including: i) relatively late age of onset and coincidence of the first symptoms with changes in the hormonal status in the organism; ii) sexual dimorphism; iii) fluctuating course and sometimes recovery; iv) parental origin effects; and v) discordance of MZ twins. We also re-analyzed several etiological theories of major psychosis from an epigenetic point of view (Petronis A, Paterson AD, Kennedy JL. Schizophrenia: an epigenetic puzzle? Schizophrenia Bulletin 25:4: 639-655, 1999; Petronis A. The genes for major psychosis: aberrant sequence or regulation? Neuropsychopharmacology, 23(1):1-12; 2000) and suggested that epigenetic mechanisms have the potential to explain a number of clinical and molecular findings that traditionally have been supporting unrelated and somewhat antagonistic theories of schizophrenia and bipolar disorder, or have not been explained at all. With regards to the field of neurobehavioral disorders the heuristic value of the epigenetic model of major psychosis lies in the possibility of integrating a wide variety of empirical data into a new theoretical framework, which provides the basis for new experimental approaches. It is important to note that epigenetic dysfunction may exhibit stability during meiosis and therefore can be transmitted from one generation to another

(Klar AJ. Propagating epigenetic states through meiosis: where Mendel's gene is more than a DNA moiety. Trends Genet 1998; 14(8):299-301; Cavalli G, Paro R. The Drosophila Fab-7 chromosomal element conveys epigenetic inheritance during mitosis and meiosis. Cell 1998; 93(4):505-18; Allen ND, Norris ML, Surani MA. Epigenetic control of transgene expression and imprinting by genotype-specific modifiers. Cell 1990 Jun 1;61(5):853-61; Silva AJ, White R. Inheritance of allelic blueprints for methylation patterns. Cell 1988 Jul 15;54(2):145-52; Morgan HD, Sutherland HG, Martin DI, and Whitelaw E (1999) Epigenetic inheritance at the agouti locus in the mouse. Nature Genetics 23: 314-8), which would simulate familial, i.e. genetic, cases of the disease.

The above description is not intended to limit the claimed invention in any manner, Furthermore, the discussed combination of features might not be absolutely necessary for the inventive solution.

The present invention will be further illustrated in the following examples. However, it is to be understood that these examples are for illustrative purposed only, and should not be used to limit the scope of the present invention in any manner.

Examples

Example 1: Detection of epigenetic abnormalities associated with schizophrenia or bipolar disorder.

Identification of the actual genes, which are epigenetically dysregulated and increase the risk to major psychosis, is not a simple task. Potentially any of the 35,000 human genes can be an epigenetic candidate for schizophrenia and bipolar disorder Based on our preliminary findings, as described below, we suggest that epigenetic analysis of multicopy DNA sequences may lead to the identification of the genes that predispose to major psychosis. At least 35% of the human genome consists of numerous copies of different transposons dispersed in the genome (NB: only ~5% of the human genome are

exons, i.e. coding sequences of functional genes) (Yoder JA, Walsh CP, Bestor TH. Cytosine methylation and the ecology of intragenomic parasites. Trends Genetics, 13(8):335-40, 1997) . The range of copies of repetitive DNA fragments varies widely: There are 106 copies of Alu sequences and 105 copies L1 elements per genome (ibid.). The general opinion is that such sequences represent excess baggage of our evolutionary heritage and do not perform any specific genomic function. This fraction of the genome is sometimes called "junk" or "parasitic" DNA. Such elements are not generally harmful to a cell as long as they do not exhibit any transcriptional activity and do not affect the integrity of the host genome. Transcriptional inactivation of the multicopy elements is achieved by their epigenetic modification. It has been widely observed that DNA methylation plays a role in silencing various types of DNA sequences. Since it is becoming evident that DNA methylation may act in concert with histone acetylation (Nan X, Campoy FJ, Bird A. MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. Cell, 88(4):471-81, 1997), chromatin conformation can also be considered a factor that plays a role in the inactivation of retrotransposons as well as any other newly integrated DNA sequence. The findings that Alu and L1 elements as well as numerous other retroelements are methylated and transcriptionally inactive in the genomes of fungi, plants, and mammals provided the basis for postulating that epigenetic DNA modification represents a host genome defense system (Bestor TH. DNA methyltransferase in genome defence. In: Epigenetic mechanisms of gene regulation. Eds: Russo VEA, Martienssen RA, Riggs AD. Cold Spring Harbor Laboratory Press, pp. 61-76, 1996; Yoder JA, Walsh CP, Bestor TH. Cytosine methylation and the ecology of intragenomic parasites. Trends Genetics, 13(8):335-40, 1997).

The epigenetic parameter may add a new dimension to the already available developments in psychiatric research. In our experiments we serendipitously detected that while the overwhelming majority of Alu sequences in the genomic DNA extracted from human brain are methylated, a small fraction of such sequences is unmethylated. The origin of such selective Alu demethylation is not clear. Without wishing to be bound by theory, this most likely represents a local failure of the epigenetic host defense system, which has no direct impact to the normal functioning of the brain. On the other hand,

such local epigenetic changes may not be limited to the Alu sequences and may extend to the surrounding genes, causing dysregulation which may be detrimental to the cells. Supporting evidence for this comes from the observation that retroelements may become demethylated because they are located in the genomic region that was subjected to genetic and epigenetic re-organization. In malignant cells, it was detected that some Alu (Rubin CM, VandeVoort CA, Teplitz RL, Schmid CW . Alu repeated DNAs are differentially methylated in primate germ cells. Nucleic Acids Research, 22(23):5121-7, 1994; Sinnett D, Richer C, Deragon JM, Labuda D. Alu RNA transcripts in human embryonal carcinoma cells. Model of post-transcriptional selection of master sequences. Journal of Molecular Biology, 226(3):689-706, 1992) and L1 (Florl AR, Franke KH, Niederacher D, Gerharz CD, Seifert HH, Schulz WA. DNA methylation and the mechanisms of CDKN2A inactivation in transitional cell carcinoma of the urinary bladder. Laboratory Investigation, 80(10):1513-22, 2000; Jurgens B, Schmitz-Drager BJ, Schulz WA. Hypomethylation of L1 LINE sequences prevailing in human urothelial carcinoma. Cancer Research, 56(24):5698-703, 1996) elements became hypomethylated and transcriptionally active.

Our working hypothesis for the experiments performed thus far is that identification of unmethylated "junk" DNA sequences in major psychosis may allow for the mapping of specific genomic regions in which epigenetic re-arrangements occurred. Dysfunction of genes that are localized in such regions may be the actual cause of psychotic symptoms, while the demethylated multicopy element sequence would serve as a reporter, a signal that allows for localization of epigenetic changes in the genome. Based on the above, we investigated hypomethylated Alu elements in major psychosis.

Results. DNA samples were extracted from the frontal cortex of 40 post-mortem brain tissues of individuals who were affected with schizophrenia and bipolar disorder as well as control individuals. In order to avoid artifacts related to partial brain DNA degradation (which may simulate hypomethylation and produce artifactual Alu amplification; see below), the following procedure was performed. Undigested total genomic DNA was fractionated on an agarose gel, the high molecular weight (>15-20kb) DNA was cut from

the gel. The gel block, containing DNA, was treated with a gel digesting enzyme, agarase. Without any additional procedures, such high quality DNA samples can be further digested with a specific restriction enzyme and subjected to further analyses. The methylation sensitive restriction enzyme, HpaII, was used for digestion of DNA and the unmethylated fraction of brain specific DNA (fragments smaller than arbitrarily selected 6kb) were separated from the methylated fraction of DNA using gel electrophoresis. The <6kb fragments were purified from the gel using glass milk. Screening for the presence of Alu's in the purified unmethylated DNA was performed using PCR and primers complementary to the Alu sequence. Alu amplicons were cloned into a vector and transformed into E.coli XL1-blue. Up to ten recombinant clones from each PCR product were sequenced from six individuals affected with major psychosis and four controls. The location of such Alu sequences were identified using human genome databases (http://genome.ucsc.edu/). It was detected that the Alu's from affected individuals in numerous cases corresponded with the genomic regions that showed evidence for linkage in genetic linkage studies of major psychosis. For example, one of the Alu sequences cloned from an affected individual mapped to chr 1q21, the region that was linked to schizophrenia (lod score of 6.5, the strongest evidence for linkage in schizophrenia genetics thus far) in large multiplex schizophrenia families (Brzustowicz LM, et al.,, 2000). In addition, an Alu clone from another psychosis patient exhibited sequence homology with 1q42, the translocation region in a schizophrenia kindred (St Clair D, et al. 1990). Other genomic regions where Alu sequences mapped to the linkage 'spots', include 5q11 (although linkage to this region [Sherrington R, et al.1988] was not replicated in other studies, two large kindreds exhibit lod scores between 2 and 3 in favor of linkage). Other identified regions include: 5q35 (chr 5 data reviewed in Crowe RR, et al. 1999), 8p23 (lod score 3.8 in a large Swedish schizophrenia kindred), 8p21, 10p14, the pericentrometric regions of chr 10 and 10q26 (Wildenauer DB, et. al. 1999), 11p15 and 11q13, 14q32 (Craddock 1999), 12p13 and 12q23-24 (Detera-Wadleigh SD. et al. 1999), and 22q13 (Numberger JI Jr, et al.1999). The 22q13 region exhibited evidence for linkage in numerous studies and harbors a deletion region in velo-cardiofacial syndrome, a disorder quite often resulting in psychotic symptoms (Chow EW, et al. 1994). For more details on the localization of the cloned Alu sequences see Figure 1. Alu sequences that

are located in the vicinity (within 100,000 bp) of coding genes are listed in Figure 2. Sequences of the cloned Alu's are provided in Figure 3.

The above results are of interest for the following reasons. First, clustering of the Alu sequences into the groups of affected individuals and controls, if replicated in an independent sample, would indicate that epigenetic changes of repetitive DNA elements in some genomic loci are specific to major psychosis. This would be a significant step forward in the light of the myriad of non-specific molecular changes in the brains of patients affected with major psychosis. Second, genomic location of the hypomethylated Alu's match with the loci that exhibit evidence for linkage to major psychosis. Traditional genetic linkage studies face major difficulties in fine mapping of the regions of susceptibility and identification of the actual gene dysfunction that leads to major psychosis. Typically the regions that exhibit evidence for linkage to major psychosis are in the range of ~10-40 cM, i.e. ~10-40 million nucleotides (Thaker GK, et al., 2001; Tsuang MT, et al. 2001; Bray NJ, and Owen MJ. 2001: Gershon ES. 2000; Nurnberger JI Jr, et al. 2000), and such regions contain hundreds of genes. Screening of such a large number of genes by traditional strategies for the detection of DNA variation is not possible. For fine mapping of prediposing genes using the transmission disequilibrium test, very large samples are required; this strategy has not been productive in psychiatric research thus far. In conclusion, the "junk" DNA-based search for major psychosis genes may represent a valuable 'shortcut' in the identification of such genes. Hypomethylated Alu's may pinpoint very specific sites of genomic DNA epigenetic dysfunction of which may cause major psychosis.

Example2: Identification of genes involved in etiology of schizophrenia or bipolar disorder based on epigenetic analysis

The genes that are located in the regions exhibiting both linkage to major psychosis and epigenetic abnormalities in Alu sequences will be subjected to a detailed analysis. Using the Celera Human Genome Database we will make a list of genes from 1q21, 5q11, 8p23, 10p14, 11p15, 12p13, 12q23-24, 22q13, chr Y, and several other loci to be further

investigated from the epigenetic point of view. It is expected that such list will include ~30 genes. We will attempt to match patients and controls for age, sex, and race. Cases with drug and alcohol abuse will not be used in the study. Treatment with neuroleptic medications is also a significant confounding factor. Neuroleptic naïve schizophrenic patients are very rare, but cases with long neuroleptic free pre-mortem intervals are quite common. For example, in a recent study, one third of brain samples were neuroleptic-free for more than 6 months (Hernandez I, et al., 2000) and during this period, ~50% of schizophrenia patients are expected to relapse (Viguera AC, et al., 1997) . If our hypothesis of epigenetic dysregulation in schizophrenia and bipolar disorder is correct, disease associated epigenetic abnormalities in the brain should recur after neuroleptic treatment is stopped. Regarding the sample size, since there are no precedents of epigenetic studies in major psychosis, power analysis on the sample size is not possible. We are starting with a relatively large sample by post-mortem brain study standards. Our plan is to investigate the prefrontal cortex from 25 post-mortem patients affected with major psychosis with >6 months of neuroleptic free period before death and a similar number of controls. Over 70 brain samples from individuals who were affected with schizophrenia or bipolar disorder as well as controls are available at our laboratory and this sample increases every year. Total mRNA from the brain tissues will be extracted using standard RNA extraction techniques (Chomczynski P,et al., 1987) and subjected to reverse transcription and quantitative PCR amplification using the Bio-Rad Real Time PCR equipment (http://www.bio-rad.com/iCycler/). This experiment will allow for the quantitative evaluation of the steady state level of the of the candidate gene. 'Is it β-actin' mRNA will serve as an internal standard for the degree of mRNA degradation. Expression of Is it β-actin is independent of the age of an individual and treatment (Schramm M, et al., 1999) and therefore can be reliably used in our experiment as an estimate of the degree of post-mortem degradation. Steady state mRNA level of each individual gene will be normalised according to its Is it β-actin mRNA data. The null hypothesis is that the group of affected individuals exhibits no differences in the steady state mRNA levels of the selected genes in comparison to the group of controls. The genes that reject the null hypothesis, i.e. the ones that exhibit statistically significant differences in steady state mRNA levels in affected tissues versus controls, will be

subjected to further analysis. The problem is that not all genes that exhibit significant differences in expression may carry epigenetic defects. Cases when changes in steady state mRNA levels that may occur within hours or even minutes after some triggers are applied, in the absence in any epigenetic changes in the genome have to be excluded. Typically, epigenetic DNA modification targets cytosines in CpG dinucleotides, each of which can be either methylated (metC) or unmethylated (C). The gold standard technique for DNA methylation analysis is based on the reaction of genomic DNA with sodium bisulfite under conditions such that cytosine is deaminated to uracil but metC remains unreacted (Frommer M, et al. 1992). Sequencing of bisulfite modified DNA reveals which cytosines were methylated and which cytosines were not. This approach has been fully operationalized in our laboratory (Popendikyte V, et al., 1999). The project can be treated as successful if a gene from the list of ~30 candidates exhibit disease specific epigenetic abnormality.

The present invention has been described with regard to preferred embodiments. However, it will be obvious to persons skilled in the art that a number of variations and modifications can be made without departing from the scope of the invention as described herein.

All references are herein incorporated by reference.

REFERENCES

. Allen ND, Norris ML, Surani MA. Epigenetic control of transgene expression and imprinting by genotype-specific modifiers. Cell 1990 Jun 1;61(5):853-61

. Bestor TH. DNA methyltransferase in genome defence. In: Epigenetic mechanisms of gene regulation. Eds: Russo VEA, Martienssen RA, Riggs AD. Cold Spring Harbor Laboratory Press, pp. 61-76, 1996.

. Bray NJ, Owen MJ. Searching for schizophrenia genes. Trends Mol Med. 2001; 7(4):169-74.

. Brzustowicz LM, Hodgkinson KA, Chow EW, Honer WG, Bassett AS.Location of a major susceptibility locus for familial schizophrenia on chromosome 1q21-q22.Science 2000 Apr 28;288(5466):678-82

. Cavalli G, Paro R. The Drosophila Fab-7 chromosomal element conveys epigenetic inheritance during mitosis and meiosis. Cell 1998; 93(4):505-18

. Chomczynski P, Sacchi N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem. 1987 Apr;162(1):156-9.

. Chow EW, Bassett AS, Weksberg R.Velo-cardio-facial syndrome and psychotic disorders: implications for psychiatric genetics. Am J Med Genet 1994;54(2):107-12

. Craddock N, Lendon C. Chromosome Workshop: chromosomes 11, 14, and 15. Am J Med Genet. 1999 Jun 18;88(3):244-54.

. Crowe RR, Vieland V. Report of the Chromosome 5 Workshop of the Sixth World Congress on Psychiatric Genetics. Am J Med Genet. 1999 Jun 18;88(3):229-32.

. Detera-Wadleigh SD. Chromosomes 12 and 16 workshop. Am J Med Genet. 1999 Jun 18; 88(3):255-9.

. Ehrlich M and Ehrlich K (1993) Effect of DNA methylation and the binding of vertebrate and plant proteins to DNA. In: Jost JP and Saluz P (eds) DNA Methylation: Molecular Biology and Biological Significance pp. 145-168. Birkhauser Verlag, Basel, Switzerland.

. Florl AR, Franke KH, Niederacher D, Gerharz CD, Seifert HH, Schulz WA. DNA methylation and the mechanisms of CDKN2A inactivation in transitional cell carcinoma of the urinary bladder. Laboratory Investigation, 80(10):1513-22, 2000.

. Frommer M, McDonald LE, Millar DS, Collis CM, Watt F, Grigg GW, Molloy PL, Paul CL. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. Proc Natl Acad Sci U S A 1992;89:1827-31.

. Gershon ES. Bipolar illness and schizophrenia as oligogenic diseases: implications for the future. Biol Psychiatry. 2000 Feb 1;47(3):240-4.

. Gonzalgo, M.L. and Jones, P.A. (1997) Mutagenic and epigenetic effects of DNA methylation. Mutat. Res. 386(2), 107-18

. Gottesman II. Schizophrenia Genesis: The Origins of Madness. New York: W.H. Freeman; 1991.

. Henikoff S, Matzke MA Exploring and explaining epigenetic effects. Trends Genet 1997;13(8):293-5

. Hernandez I, Sokolov BP. Abnormalities in 5-HT2A receptor mRNA expression in frontal cortex of chronic elderly schizophrenics with varying histories of neuroleptic treatment. J Neurosci Res. 2000; 59(2):218-25.

. Jones PL, Veenstra GJ, Wade PA, Vermaak D, Kass SU, Landsberger N, Strouboulis J, and Wolffe AP (1998) Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nature Genetics 19: 187-91.

. Jurgens B, Schmitz-Drager BJ, Schulz WA. Hypomethylation of L1 LINE sequences prevailing in human urothelial carcinoma. Cancer Research, 56(24):5698-703, 1996.

. Klar AJ. Propagating epigenetic states through meiosis: where Mendel's gene is more than a DNA moiety. Trends Genet 1998; 14(8):299-301.

. Lyko, F. and Paro, R. (1999) Chromosomal elements conferring epigenetic inheritance. Bioessays 21(10), 824-32

. Morgan HD, Sutherland HG, Martin DI, and Whitelaw E (1999) Epigenetic inheritance at the agouti locus in the mouse. Nature Genetics 23: 314-8.

. Nan X, Campoy FJ, Bird A. MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. Cell, 88(4):471-81, 1997.

. Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN, and Bird A (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature 393: 386-9.

. Numberger JI Jr, Foroud T. Chromosome 6 workshop report. Am J Med Genet. 1999 Jun 18;88(3):233-8.

. Numberger JI Jr, Foroud T. Genetics of bipolar affective disorder. Curr Psychiatry Rep 2000 Apr;2(2):147-57

. Petronis A. Human morbid genetics revisited: relevance of epigenetics. Trends Genet 2001 Mar;17(3):142-6.

. Petronis A, Paterson AD, Kennedy JL. Schizophrenia: an epigenetic puzzle? Schizophrenia Bulletin 25:4: 639-655, 1999

. Petronis A. The genes for major psychosis: aberrant sequence or regulation? Neuropsychopharmacology, 23(1):1-12; 2000.

. Popendikyte V, Laurinavicius A, Paterson AD, Macciardi F, Kennedy JL, Petronis A. DNA methylation at the putative promoter region of the human dopamine D2 receptor gene. Neuroreport 1999;10:1249-55.

. Razin, A. and Shemer, R. (1999) Epigenetic control of gene expression. Results Probl. Cell. Differ. 25, 189-204

. Riggs A, Xiong Z, Wang L, and LeBon JM (1998) Methylation dynamics, epigenetic fidelity and X chromosome structure. In: Wolffe AP (ed) Epigenetics, pp. 214-227. John Wiley & Sons, Chistester.

. Robertson KD and Wolffe AP (2000) DNA methylation in health and disease. Nature Review Genet 1:11-9

. Rubin CM, VandeVoort CA, Teplitz RL, Schmid CW . Alu repeated DNAs are differentially methylated in primate germ cells. Nucleic Acids Research, 22(23):5121-7, 1994.

. Schramm M, Falkai P, Tepest R, Schneider-Axmann T, Przkora R, Waha A, Pietsch T, Bonte W, Bayer TA. Stability of RNA transcripts in post-mortem psychiatric brains. J Neural Transm. 1999;106(3-4):329-35.

. Sherrington R, Brynjolfsson J, Petursson H, Potter M, Dudleston K, Barraclough B, Wasmuth J, Dobbs M, Gurling H. Localization of a susceptibility locus for schizophrenia on chromosome 5. Nature. 1988 Nov 10;336(6195):164-7.

. Siegfried Z, Eden S, Mendelsohn M, Feng X, Tsuberi BZ, Cedar H. DNA methylation represses transcription in vivo. Nat Genet 1999 Jun;22(2):203-206

. Silva AJ, White R. Inheritance of allelic blueprints for methylation patterns. Cell 1988 Jul 15;54(2):145-52

. Sinnett D, Richer C, Deragon JM, Labuda D. Alu RNA transcripts in human embryonal carcinoma cells. Model of post-transcriptional selection of master sequences. Journal of Molecular Biology, 226(3):689-706, 1992.

. St Clair D, Blackwood D, Muir W, Carothers A, Walker M, Spowart G, Gosden C, Evans HJ. Association within a family of a balanced autosomal translocation with major mental illness. Lancet. 1990 Jul 7;336(8706):13-6.

. Thaker GK, Carpenter WT Jr. Advances in schizophrenia. Nat Med. 2001 Jun;7(6):667-71.

. Tsuang MT, Stone WS, Faraone SV. Genes, environment and schizophrenia. Br J Psychiatry Supl. 2001 Apr;40:s18-24.

. Viguera AC, Baldessarini RJ, Hegarty JD, van Kammen DP, Tohen M. Clinical risk following abrupt and gradual withdrawal of maintenance neuroleptic treatment. Arch Gen Psychiatry 1997; 54(1):49-55

. Wildenauer DB, Schwab SG. Chromosomes 8 and 10 workshop. Am J Med Genet. 1999 Jun 18;88(3):239-43.

. Xu G-L, Bestor TH. Nature Genetics 17: 376-379, 1997.

. Yoder JA, Walsh CP, Bestor TH. Cytosine methylation and the ecology of intragenomic parasites. Trends Genetics, 13(8):335-40, 1997.

THE EMBODIMENTS OF THE INVENTION IN WHICH AN EXCLUSIVE PROPERTY OF PRIVILEGE IS CLAIMED ARE DEFINED AS FOLLOWS:

1. Method of detecting an epigenetic abnormality associated with a complex non-Mendelian disease, said method comprising:

a) extraction of genomic DNA from a sample that exhibits characteristics of a non-Mendelian disease;

b) digestion of said genomic DNA with a methylation-sensitive restriction enzyme to produce a pool of restricted DNA fragments;

c) fractionation of said pool of restricted DNA fragments to obtain DNA fragments of a desired size;

d) amplification of at least a segment of said DNA fragments of a desired size with primers that anneal to an endogenous DNA element to produce a PCR product;

e) cloning of said PCR product into a sequencing vector;

f) sequence determination of said PCR product to obtain a sequence of said PCR product;

g) comparing said sequence against a genomic database to assign a locus for said epigenetic abnormality associated with a non-Mendelian disease.

2. The method of claim 1, wherein said non-Mendelian disease is selected from the group consisting of schizophrenia, bipolar disorder, cancer, and diabetes.

3. The method of claim 1, wherein said sample that exhibits characteristics of a non-Mendelian disease is brain tissue.

4. The method of claim 3, wherein said sample that exhibits characteristics of a non-Mendelian disease is selected from the group consisting of frontal cortex and prefrontal cortex.

5. The method of claim 1, wherein said desired size is less than 10 kb.

6. The method of claim 1, wherein said endogenous DNA element is a multi-copy DNA element.

7. The method of claim 6, wherein said multi-copy DNA element is selected from the group consisting of LINE, SINE, L1, and Alu..

8. The method of claim 1, wherein said methylation-sensitive restriction enzyme is selected from the group consisting of AatII (GACGTC); Bsh1236I (CGCG); Bsh1285I (CGRYCG); BshTI (ACCGGT); Bsp68I (TCGCGA); Bsp119I (TTCGAA); Bsp143II (RGCGCY); Bsu15I (ATCGAT); Cfr10I (RCCGGY); Cfr42I (CCGCGG); CpoI (CGGWCCG); Eco47III (AGCGCT); Eco52I (CGGCCG); Eco72I (CACGTG); Eco105I (TACGTA); EheI (GGCGCC); Esp3I (CGTCTC); FspAI (RTGCGCAY); Hin1I (GRCGYC); Hin6I (GCGC); HpaII (CCGG); Kpn2I (TCCGGA); MluI (ACGCGT); NotI (GCGGCCGC); NsbI (TGCGCA); PauI (GCGCGC); PdiI (GCCGGC); Pfl23II (CGTACG); Psp1406I (AACGTT); PvuI (CGATCG); SalI (GTCGAC); SmaI (CCCGGG); SmuI (CCCGC); TaiI (ACGT); and TauI (GCSGC).

9. Method of identifying a gene having an epigenetically altered expression pattern that contributes to a non-Mendelian disease in an organism, said method comprising:

a) extraction of genomic DNA from a sample that exhibits characteristics of a non-Mendelian disease;

b) digestion of said genomic DNA with a methylation-sensitive restriction enzyme to produce a pool of restricted DNA fragments;

c) fractionation of said pool of restricted DNA fragments to obtain DNA fragments of a desired size;

d) amplification of at least a segment of said DNA fragments of a desired size with primers that anneal to an endogenous DNA element to produce a PCR product;

e) cloning of said PCR product into a sequencing vector;

f) sequence determination of said PCR product to obtain a sequence of said PCR product;

g) comparing said sequence against a genomic database to assign a locus for said epigenetic abnormality associated with a non-Mendelian disease;

h) searching said database to identify a gene located proximal to said locus;

i) comparing expression patterns of said gene located proximal to said locus within a test sample that exhibits characteristics of said non-Mendelian disease with expression patterns of a corresponding gene within a control sample to identify said gene having an epigenetically altered expression pattern.

10. A gene isolated by the method of claim 9.

11. Method of isolating a probe for detecting an epigenetic abnormality associated with a non-Mendelian disease, said method comprising:

a) extraction of genomic DNA from a sample that exhibits characteristics of a non-Mendelian disease;

b) digestion of said genomic DNA with a methylation-sensitive restriction enzyme to produce a pool of restricted DNA fragments;

c) fractionation of said pool of restricted DNA fragments to obtain DNA fragments of a desired size;

d) amplification of at least a segment of said DNA fragments of a desired size with primers that anneal to an endogenous DNA element to produce a PCR product;

e) using said PCR product as said probe to detect said epigenetic abnormality associated with a non-Mendelian disease in another sample.

12. A probe isolated by the method of claim 11.

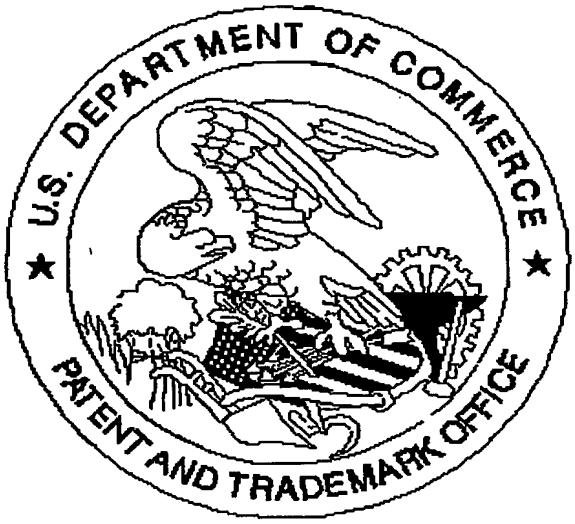
13. The method of of claim 1 wherein said detecting an epigenetic abnormality associated with a complex non-Mendelian disease, is used to diagnose an epigenetic abnormality associated with a complex non-Mendelian disease.

ABSTRACT OF THE DISCLOSURE

The invention can be summarized as follows. A method comprising extraction of genomic DNA from a diseased tissue or diseased population of cells; hydrolysis of this DNA with methylation-sensitive restriction enzymes, and subsequent fractionation of DNA fragments and purification of DNA fragments of a desired size, PCR amplification using primers that hybridize to endogenous DNA elements including. PCR products of such elements are cloned and sequenced using standard molecular biology techniques known to the skilled artisan and the resultant sequences are mapped on the genome using any commercially or publicly available human genome database. These cloned endogenous DNA elements indicate a loci of putative epigenetic abnormality or epigenetic dys-regulation and indicates genes that predispose a patient to a complex, non-Mendelian, multi-factorial disease, such as, but not limited to, cancers, diabetes, schizophrenia, or bipolar disorder.

United States Patent & Trademark Office

Office of Initial Patent Examination -- Scanning Division



Application deficiencies found during scanning:

☐ Page(s)_____ of _____ were not present for scanning. (Document title)

☐ Page(s)_____ of _____ were not present for scanning. (Document title)

☑ ***Scanned copy is best available.*** There are 65 page of drawings. Not 66

ATTACHMENT 1. LOCALIZATION OF ALU SEQUENCES THAT MATCH TO THE GENOMIC REGIONS THAT EXHIBITED EVIDENCE FOR LINKAGE TO MAJOR PSYCHOSIS

SZ - Alu clones from individuals affected with schizophenia
BD - Alu clones from individuals affected with bipolar disorder
MD - major depression
CTRL - control samples

Sample Name (matched bp, %, chr band) numer of ccgg sites	Homology length in bp; %	Chromosomal location	Evidence for linkage or association to schizophrenia or bipolar disorder
SZe-32m56	189, 99.5 %	6p22.3	Eckstein GN, Schwab SG, Maier W, Wildenauer DB. 1998. Searching for candidate genes for schizophrenia in chromosome 6p22,23: isolation of a BAC contig spanning 3.5 megabases. Am J Med Genet 81.530.
Sch37-9RR E-283m56SZ	160, 98.2 % 190, 99.5%	10p14 10p14	10p11-15 Faraone et al. (1998) nonparametric LOD scores at markers D10S1423 and D10S582 were 3.4 (P = .0004) and 3.2 (P = .0006), respectively. Schwab et al. (1998a), [1]nonparametric LOD score of 3.2 (P = .0007) at marker D10S1714(Schwab et al. 1998) (Straub et al. 1998)Straub et al. (1998) LOD score of 1.91 (P = .006) at with markers D10S1426 and D10S674
SZr-37m56	183, 96.5 %	11q14.2	Mulcrone J, Whatley SA, Marchbanks R, Wildenauer D, Altmark D,

[1] Schwab SG, Hallmayer J, Albus M, Lerer B, Hunses C, Kanyas K, Segman R, Borrman M, Dreikorn B, Lichtermann D, Rietschel M, Trixler M, Maier W, Wildenauer DB 1998 Further evidence for a susceptibility locus on chromosome 10p14-p11 in 72 families with Karayiorgou M, Kasch L, Lasseter VK, Hwang J, Elango R, Bernardini DJ, Kamberland M, Tobb R, Francomano CA, Wolyniec PS, et al. (1994) Report from the Maryland Epidemiology Schizophrenia Linkage Study: no evidence for linkage between schizophrenia and a number of candidate and other genomic regions using a complex dominant model *Am J Med Genet* 54:145-57 @ by nonparametric linkage analysis. Am J Med Genet 81 302–307.

Figure 1

			Daoud H, Gur E, Ebstein RP. Lerer B 1995. genetic linkage analysis of schizophrenia using chromosome 11q13–24 markers in Israeli pedigrees Am J Med Genet 60:103–108.
E-318_m74_SZ	206, 97 7 %	22q12.2	22q11-13, Pulver et al. (1994*a*)(Pulver et al. 1994a; Pulver et al. 1994b; Pulver et al. 1994c) LOD score of 2.82 at marker locus IL2RB; (*P* = 009) The implicated region is near the velocardiofacial syndrome(VCFS) deletion, Lasseter et al. 1995(Lasseter et al. 1995) Polymeropoulos (Polymeropoulos et al. 1994)et al. 1994 Coon (Coon et al. 1994a. Coon et al. 1994b)et al 1994*a* Stober (Stober et al. 2000)et al 2000 Myles-Worsley(Myles-Worsley et al. 1999) et al. 1999
E-305_m740_SZ E-221_m37_SZ E-267_m50_Ctrl E-288_m56_SZ E-289_m56_SZ E-297_m740_SZ E-295_m740_SZ E-294_m740_SZ E-293_m56_SZ E-286_m56_SZ E-252_m48_SZ E-244_m48_SZ E-130_m37_SZ SZm74-E-59 SZm74-E-58 SZm74-E-50 SZb-_M37-(	191, 100 %	Yq12, Yq11 23, Yq11.223	Yq11.23 and Yq12(Alitalo et al 1988) Alitalo T, Tiihonen J, Hakola P, de la Chapelle A. 1988

Figure 1 Continued

SZb-_M37-7 SZC-_M37-5 SZC-_M37-2 SZC-_M37-26 SZC-_M37-15 SZC-_M37-7 SZC-_M37-5 SZD-_M37-14 SZRevCom48_E-33 SZRevCom48_E-39 SZm37-E-13_m37-7 Sch37-1 Sch37-6 Sch37-7 E-284m56SZ			
E-267_m50_Ctrl E-261_m50_Ctrl E-167m50Ctrl E-275m50Ctrl E-281m50Ctrl RevE-270m50Ctrl	191. 100 %	Yq12, Yq11 23, Yq11.223	Yq11.23 and Yq12(Alitalo et al. 1988) Alitalo T. Tiihonen J, Hakola P. de la Chapelle A. 1988
CONTROLS			
Ctrlm57-E-6	187; 99%	1q31.1	D1S2141 1q32-q41 Hovatta et al. (1998) (Hovatta et al. 1998) *1q32-41* Hovatta et al. (1999) (Hovatta et al. 1999) LOD score of 3.82 at marker D1S2891
RevE-169m50Ctrl	179; 94.8%	1q31.1	
E-271m50Ctrl	155, 90.6 %	1q32.1	

Figure 1 Continued

			Schizophrenia Hovatta et al. (1998) (Hovatta et al. 1998) D1S2141 1q32-q41 Lod score 90% penetrance Lod score = 3.73
Ctrlm50E-49	185, 98 %	2q35	Event-related brain potential P3 Almasy et al. (1998)(Almasy and Blangero 1998) Between D2S425 and D2S434 2q33-q37 Bivariate quantitative linkage analysis Lod score = 3.28
Ctrlm57-E-3	191, 100 % or 189, 99.5 %	5q33.2 18q22.2	*5q22-31* 5q31 LOD score of 3.35 (P = 0002) at marker D5S804 5q23.3 Straub et al. (1997) (Straub et al. 1997) Marker D5S399 at 5q31 5q31.3-q35.1 was presented by Shink et al [1998][2](Morissette et al. 1999) Shink E. Morissette J, Rochette D, Bordeleau L. Plante M, Villeneuve A, Barden N. 1998. Bipolar affective disorder susceptibility loci on chromosomes 5 and 21: heterogeneity in a homogeneous population in Quebec.

[2] Shink E, Morissette J, Rochette D Bordeleau L Plante M Villeneuve A Barden N 1998 Bipolar affective disorder susceptibility loci on chromosomes 5 and 21 heterogeneity in a homogeneous population in Quebec Am J Med Genet 81(6):541-542

199 BIPOLAR AFFECTIVE DISORDER SUSCEPTIBILITY LOCI ON CHROMOSOMES 5 AND 21: HETEROGENEITY IN A HOMOGENEOUS POPULATION IN QUEBEC. Shink E.,1 Morissette J,1 Rochette D,2 Bordeleau L 1 Plante M,1 Villeneuve A,3 and Barden N.1 1Neuroscience, CHUL, Quebec, Canada G1V 4G2. 2Complexe Hospitalier de la Sagamie, Chicoutimi, Québec. 3Clinique Roy-Rousseau, Québec.
Following a linkage study that yielded strong evidence for a susceptibility locus for bipolar affective disorder (BP) on chromosome 12 it appeared that one branch of an extremely large pedigree from the Saguenay-Lac-St-Jean region of Québec showed strongest linkage for regions on chromosomes 5 and 21 Two-point linkage analysis was carried out in this branch using a model of dominant transmission and with diagnoses of BP I, recurrent unipolar and BP II considered as affected, those with single episode major depression classified as unknown and all other diagnoses as unaffected LOD scores greater than 1 for four successive loci: D5S673 (1 42), D5S410 (1 20), D5S412 (1 87) and D5S422 (1 82) and for two

Figure 1 Continued

Ctrlm57-E-5 .	186. 97.4 %	13q14.11	13q14-32. Blouin et al. (1998)(Blouin et al. 1998) nonparametric LOD score of 4.18 (P = 00002). near D13S174 on 13q32 Brzustowicz et al. (1999)
E-166m50Ctrl	181. 100 %	18q23	[2]Ewald et al [1998] found increased haplotype sharing with distal markers at 18q23 in eight BPI patients from the Faroe Islands, in a region also suggested by Freimer et al. [1996].
E-279m50Ctrl	132, 94.7 %	18p11 23	18p11 2 and 18q12.1-q12.3 for BP and SZ., [3]Gershon et al [1998]

other successive loci· D21S265 (1.68) and D21S263 (1 13) were obtained. Simulation on this nucleus shown that if theta is greater than 0 3, the probability to get a LOD score greater than 1.6 is less than 0 03 The NPL scores calculated with the Genehunter program demonstrated strong increased allele sharing in the neighbourhood of D5S673 - D5S410 and D21S265-D21S263 with, respectively, maximum scores of 10 51 and 11 59 (significant at the 0.004 level) For loci D5S673 D5S410, D5S412, and D5S422, all affected persons, with one exception, share the same haplotype. A common haplotype at D21S265 and D21S263 is shared by all affected persons and individuals classified as unknown have only one of the two common haplotypes. Simulation, nonparametric, and haplotype analysis thus suggest the presence of additional susceptibility loci on chromosomes 5 and 21 in this population.

[2] Ewald H, Nyegaard M, Wang A, Vang M, Mors O, and Kruse TA. 1998 A search for a shared segment of chromosome 18 in patients with bipolar affective disorder from the Faeroe Islands Am J Med Genet 81(6) 441.

197. A SEARCH FOR A SHARED SEGMENT OF CHROMOSOME 18 IN PATIENTS WITH BIPOLAR AFFECTIVE DISORDER FROM THE FAEROE ISLANDS
Ewald H, [1]Nyegaard M, [2]Wang A, [3]Vang M, [4]Mors O, and Kruse TA. [1]Institute for Basic Psychiatric Research Departments of [2]Psychiatric Demography and [3]Biological Psychiatry, Psychiatric Hospital, Aarhus Denmark, [4]Department of Psychiatry, National Hospital Torshavn, Faeroe Islands, [5]Department of Psychiatry, Municipal Hospital, Copenhagen University Hospital, Denmark, [6]Institute of Human Genetics Aarhus University Denmark.
The Faeroe Islands were populated at the same time as Iceland i.e. around 1,100 years ago The size of the population has recently increased by more than 10-fold to around 45,000 today This recent increase has mainly been by reproduction and together with the relative isolated geographical location thus makes the Faeroese population ideal for mapping of disease genes. Chromosome 18 has probably been the most thoroughly searched chromosome in bipolar affective patients At least three regions on the long arm and one on the short arm have received support from more than one research group The present study searched for segment sharing on chromosome 18 among distantly related lithium responding bipolar patients and controlled from the same internal subisolate of the Faeroe Islands using more than 30 microsatellite markers

[3] Gejman PV Gershon ES, Zhang F Badner JA, Sanders AR, Ceo Q, et al 1998 Progress of genome scan of the NIMH intramural schizophrenia collection Am J Med Genet 81(6) 443
PROGRESS OF GENOME SCAN OF THE NIMH INTRAMURAL SCHIZOPHRENIA COLLECTION P V Gejman, E S Gershon, J Zhang, J.A. Badner, A.R Sanders, Q. Ceo, C Markey, and L R Goldin. Unit on Molecular Clinical Investigation, Clinical Neurogenetics Branch, NIMH Bethesda, MD 20892-1274
Objective: To perform a genomic scan and replication mapping in 69 families with one or more affected sib pair (ASP) (all ASPs 95, independent ASPs: 73) with schizophrenia or schizoaffective disorder. Methods: We used the Human Screening Set/Weber Version 6 2 for the systematic mapping and we analyzed the data with allele sharing identical by descent (IBD) and multipoint maximum likelihood scores (MLS) statistics We followed up with high resolution mapping of chromosomal regions from our genome scan with a P value ¯0.01 or suggestive or significant linkages from other schizophrenia data sets Results: We have previously reported evidence for linkage to schizophrenia with chromosome 6q markers

Figure 1 Continued

			WCPG High density screen chromosome 18; average density 3.25 cM BP: 22 multiplex BP families [see (Berrettini et al. 1994)Berrettini et al. 1994] c ASM I: BPI, BPII, SA c ASM II ASM I + RUP c Nonparametric analysis (ASPEX) c ASMI. highest peak on 18p11.2 (lod 4.2 32; p 4 0.00054) c ASMII: smaller peak closer to 18ptel (lod 1.44; p 4 0.005) c Smaller peak at 18q21 (lod [illegible] not significant) c Confirmation previous evidence for linkage to 18p11.2
Ctrlm57-E-4.	193, 100 %	22q12.2	22q11-13. Pulver et al. (1994*a*)(Pulver et al. 1994a; Pulver et al. 1994b; Pulver et al 1994c) LOD score of 2.82 at marker locus IL2RB same general region (P = .009) The implicated region is near the velocardiofacial syndrome(VCFS) deletion, Lasseter et al. 1995(Lasseter et al. 1995) Polymeropoulos (Polymeropoulos et al. 1994)et al. 1994 Coon (Coon et al. 1994a: Coon et al 1994b)et al. 1994*a* Stober (Stober et al. 2000)et al. 2000 Myles-Worsley(Myles-Worsley et al. 1999) et al. 1999
Ctrlm57-6-E-1	155, 87 5 %	22q13.2	*22q11-13* Baron(Baron 1990; Baron 1995) 1990, 1995, Baron et al (Baron et al. 1990). 1990; Risch (Risch 1990a; Risch 1990b)1990*a*; Pauls (Pauls 1993)1993. Spence (Spence et al. 1993)et al. 1993; Cloninger (Cloninger 1994) 1994; Lander and Kruglyak 1995(Lander

from three data sets We have also performed high resolution mapping of several candidate areas in chromosomes 5 10, 13. 15. and 18 in the genome scan data set The most positive evidence with two point non-parametric analyses was detected in 10p (D10S189: 2.54—ASPEX. MLOD. independent ASPs. and 3 39, all ASPs) These scores decreased. however. in multipoint analysis (MLS < 1 5) In 15q positive LOD scores were obtained (D15S659 1 69—ASPEX. MLOD. independent ASPs; and 1 04, all ASPs) No evidence for linkage was detected in chromosomes 5. 13. and 18 Conclusions: No evidence for linkage was found in chromosomes 5. 13, and 18 Modestly supportive evidence for linkage was detected in chromosomes 10p and 15q

Figure 1 Continued

60[illegible]6818 .060602

			and Kruglyak 1995): Owen and Craddock (Owen and Craddock 1996) 1996).

BD43-15	190, 98.7 %	21q21 3	C21q21-22 Susceptibility Locus for Bipolar and Unipolar Affective Disorders Repeated From Gurling [1998](Gurling 1998).
BD43-6	190, 99%	1q21.1	*1q21-22* Brzustowicz et al (2000)(Brzustowicz et al. 2000; Maziade et al. 2002) heterogeneity LOD score of 6 50 was found between markers D1S1653 and D1S1679. Shaw et al. 1998(Shaw et al. 1998) 1q21 Dror et al. 1999(Dror et al. 1999) A potassium-channel gene (Hkca3/KCNN3) mapped to 1q21 - Austin et al 1999).(- hKCa3/KCNN3) (Austin et al. 1999)
RevE-77m43BD	191, 99.5 %	1p31 1	Bipolar disorder Rice et al. (1997)[5] D1S1648 1p31-p21 Sib-pair analysis MLOD2 5
BDd_M34-14BD (	187. 99 %	2p23.2).	Schizophrenia Blouin et al. (1998) (Blouin et al 1998) D2S405 2p22.1 Nonparametric lod score NPL = 1.26 (p = 0.104)
E-79m43BD	186. 96.9 %	2q37 3	Event-related brain potential P3 Almasy et al. (1998)(Almasy and Blangero 1998) Between D2S425 and D2S434 2q33-q37 Bivariate quantitative linkage analysis Lod score = 3 28
E-78m43BD E-83m43BD	192, 100 % 192, 100 % 192, 100 %	5q13.2; 5q22.2; 5q13.3;	5q11-13 Sherrington[6] et al. (1988)(Sherrington et al. 1988a; Sherrington et al. 1988b). British and Icelandic pedigrees(a LOD score of 6.49, under a dominant model Maximum LOD score of 4.37at locus D5S111

[5] Rice JP (1997) The role of meta-analysis in linkage studies of complex traits Am J Med Genet 74 112-114

Figure 1 Continued

	192, 100 %	16q23 1	5q11-13 Silverman[7] et al. (1996)(Silverman et al 1996) (Straub et al. 1997). (Bennett et al. 1997) Straub RE. MacLean CJ, O'Neill FA, Walsh D, Kendler KS. 1997 Support for a possible schizophrenia vulnerability locus in region 5q22-31 in Irish families Mol Psychiatry 2.148-155. Bennett RL. Karayiorgou M, Sobin CA, Norwood TH, Kay MA. 1997. Am J Hum Genet 61:1450-1454.
BDd_M34-19BD.	192, 100 %	10p14 or 10p13	10p11-15 Faraone et al. (1998) nonparametric LOD scores at markers D10S1423 and D10S582 were 3.4 (P = .0004) and 3 2 (P = .0006). respectively. Schwab et al. (1998a). [9]nonparametric LOD score of 3.2 (P = .0007) at marker D10S1714(Schwab et al. 1998) (Straub et al. 1998)Straub et al. (1998) LOD score of 1 91 (P = .006) at with markers D10S1426 and D10S674
E-62m34BD	192, 100 %	10p14	10p11-15 Faraone et al. (1998) nonparametric LOD scores at markers D10S1423 and D10S582 were 3.4 (P = 0004) and 3.2 (P = .0006),

[6] Sherrington R. Brynjolfsson J. Petursson H. Potter M, Dudleston K. Barraclough B, Wasmuth J, Dobbs M, Gurling H (1988) Localization of a susceptivility locus for schizophrenia on chromosome 5. Nature 336:164-167 First citation in article | PubMed

[7] Kalsi G, Mankoo B, Curtis D Sherrington R, Meimer G, Brynjolfsson J Sigmundsson T, Read T, Murphy P, Petursson H, Gurling H (1999) New DNA markers with increased informativeness show diminished support for a chromosome 5q11-13 schizophrenia susceptibility locus and exclude linkage in two new cohorts of British and Icelandic families. Ann Hum Genet 63.235-247 First citation in article | PubMed

[9] Schwab SG Hallmayer J Albus M Lerer B, Hanses C Kanyas K Segman R, Borrmann M, Dreikorn B Lichtermann D, Rietschel M, Trixler M, Maier W, Wildenauer DB 1998 Further evidence for a susceptibility locus on chromosome 10p11-p15 in 72 families with schizophrenia by nonparametric linkage analysis Am J Med Genet 81 102-107

Figure 1 Continued

			respectively. Schwab et al. (1998a), [9]nonparametric LOD score of 3.2 (P =.0007) at marker D10S1714(Schwab et al. 1998) (Straub et al 1998)Straub et al. (1998) LOD score of 1 91 (P =.006) at with markers D10S1426 and D10S674
BDC-_M34-10BD BDC-_M34-1BD BD34-5 BD34-8 BD43-1 BD43-2	191, 100 %	Yq12, Yq11 23, Yq11.223	Yq11.23 and Yq12(Alitalo et al 1988) Alitalo T, Tiihonen J, Hakola P, de la Chapelle A. 1988
MDC-_M39-2 MDD-_M39-14 MD39-4 MD39-6 MD39-8 MD39-10 E-66m39MD	191. 100 %	Yq12, Yq11.23, Yq11 223	Yq11.23 and Yq12(Alitalo et al. 1988) Alitalo T, Tiihonen J, Hakola P, de la Chapelle A. 1988

[9] Schwab SG, Hallmayer J, Albus M, Lerer B, Hanses C, Kanyas K, Segman R, Borrmann M, Dreikorn B, Lichtermann D, Rietschel M, Trixler M, Maier W, Wildenauer DB 1998 Further evidence for a susceptibility locus on chromosome 10p14-p11 in 72 families with Karayiorgou M, Kasch L, Lasseter VK, Hwang J, Elango R, Bernardini DJ, Kimberland M, Babb R, Francomano CA, Wolyniec PS, et al. (1994) Report from the Maryland Epidemiology Schizophrenia Linkage Study: no evidence for linkage between schizophrenia and a number of candidate and other genomic regions using a complex dominant model *Am J Med Genet* 54:345-53 by nonparametric linkage analysis Am J Med Genet 81 302-307

Figure 1 Continued

ATTACHMENT 2. GENES LOCATED IN THE CLOSE VICINITY TO THE CLONED *ALU* SEQUENCES

SZ - Alu clones from individuals affected with schizophenia
BD - Alu clones from individuals affected with bipolar disorder
MD - major depression
CTRL - control samples

References in the brackets in the right hand side column indicate the papers in which implication of the detected genes in major psychosis was discussed.

Clone Name	Homology length in bp; %	Chromosomal location	Genes located in the close vicinity (within 100,000 bp)
E-285_m56_SZ E-290_m56_SZ	198; 99.5% 189; 99.5%	1q31.1 1q31.1	prostaglandin-endoperoxide synthase 2, PTGS2 {Das, 1998 #1;Smythies, 1997 #2;Geling, 1991 #3}
E-149_m48_SZ	197; 99.5%	1q42.3	ryanodine receptor 2 (cardiac), RYR2
E-154_m56_SZ	188; 99%	2q33.1	general transcription factor IIIC, polypeptide 3, GTF3C3
SZcRev_M37-6	187; 99%	5q14.1	MSH3, mutS (E. coli) homolog 3 CENPH, kinetochore protein CENP-H CFDP1, craniofacial development protein 1 (Goodman, 1996 #4) IL1A, interleukin 1, alpha CRHBP, corticotropin releasing hormone-binding protein
SZe-32m56	189, 99.5 %	6p22.3	Ataxin 1, SCA1 6 papers found on Schizophrenia. 3 items found on bipolar {Culjkovic, 2000 #100;Li, 1999 #101;Joo, 1999 #102;Pujana, 1997 #103;Morris-Rosendahl, 1997 #104,Wang, 1996 #105} {Morris-Rosendahl, 1997 #40;Fernandez Piqueras, 1995 #41}
E-311_m74_SZ	201, 100 %	8p21.3	docking protein 2, 56kD, DOK2
SZe-35m56	189, 99.5 %	8q24.23	hypothetical protein FLJ10901, FLJ10901
E-322_m74_SZ	192, 100%	7p22.3	C4S-2, chondroitin 4-O-sulfotransferase 2 EIF3S9, eukaryotic translation initiation factor 3
SZm74-E-60.	186, 99.5 %	8p23.1	hypothetical protein MGC16279

Figure 2

SZe-37m56	183, 96.5 %	11q14.2	embryonic ectoderm development, EED
E-310_m74_SZ	192. 100 %	14q21.3	ribosomal protein S29. RPS29 {Gentry. 2000 #49,Watanabe. 1996 #50} {Watanabe, 1994 #106}
E-313_m74_SZ	207, 97.7 %	15q26.3	MADS box transcription enhancer factor 2., MEF2A {Turner. 1997 #109}
E-258_m48_SZ	199, 98.6 %	17q21.33	distal-less homeobox 4, DLX4
E-16_m37_SZ	191, 99.5%	17q23.2	tousled-like kinase 2, TLK2
E-319_m74_SZ	196, 100 %	18p11.32	Hypothetical protein FLJ23017, FLJ23017 highly expressed in cancer, rich in leucine, HEC
E-315_m74_SZ E-321_m74_SZ E-315_m74_SZ	191. 100 %	19q12	ubiquinol-cytochrome c reductase, Rieske, UQCRFS1 {Johnston-Wilson. 2000 #53}
E-315_m74_SZ E-321_m74_SZ E-315_m74_SZ	191, 100 %	19p13.2	hypothetical protein FLJ14356, FLJ14356 gonadotropin inducible transcription, GIOT-2 Kruppel-type zinc finger (C2H2). ZK1
E-251_m48_SZ E- 2531_m48_SZ E- 2532_m48_SZ	198, 99.5 % 189, 100% 188, 98.5%	19p13.11 19p13.11 19p13.11	hypothetical protein FLJ13659, FLJ13659
E-325_m74_SZ	204. 96.7 %	19p13.11	hypothetical protein FLJ13659
E-178_m74_SZ	205, 98.1 %	19q13 12	zinc finger protein HZF10, ZNF345 Takase. 2001 #54;Ogura, 2001 #55,Sun. 2001 #56
E-246_m48_SZ	192, 100 %	20p12 3	hypothetical protein MGC4816, MGC4816
SZd_M37-3.	190, 100 %	20q13.2	LOC57167, similar to SALL1 (sal (Drosophila)-like
SZd_M37-10.	190, 97 9 %	20q13.2	LOC57167. similar to SALL1 (sal (Drosophila)-like
E-318_m74_SZ	206. 97.7 %	22q12.2	oncostatin M, OSM
E- 305_m740_SZ E-221_m37_SZ	191, 100 %	Yq12, Yq11.23. Yq11 223	variable charge, Y chromosome, 2 protein. VCY2

Figure 2 Continued

Figure 2 Continued

Sch37-6 Sch37-7 E-284m56SZ			
E-312_m74_SZ	172. 96 1 %	Yq12. Yq11.23, Yq11 223	variable charge, Y chromosome, 2 protein. VCY2
Ctrlm57-E-6 RevE- 169m50Ctrl	187, 99% 179. 94.8%	1q31.1 1q31.1	LOC51235, hypothetical protein PTGS2, prostaglandin-endoperoxide synthase 2 (Das, 1998 #1;Smythies. 1997 #2;Gelmg, 1991 #3} PIN1L, protein (peptidyl-prolyl cis/trans isomerase)
Ctrlm50E-49	185: 98%	2q35	long-chain fatty-acid-Coenzyme A ligase 3, FACL3
RevE- 119m57Ctrl	192: 99 1% 181; 97.4%	3p22.2 3p22.1	SEC22C, vesicle trafficking protein, isoform a
Ctrlm57-E-3	191; 100% or 189, 99.5%	5q33 2 18q22.2	MRPL22, mitochondrial ribosomal protein L22 C5orf4. putative tumor suppressor PTGER4, prostaglandin E receptor 4 (subtype EP4) (Yeragani, 1987 #5}
Ctrl_m50-26	73. 86 2 %	8q11.23	lysophospholipase I, LYPLA1
gDNA Ctrl	190, 99 5%	10p14	CUG triplet repeat, RNA-binding protein 2, CUGBP2 GATA-binding protein 3. GATA3
gDNA Ctrl	187, 100 %	10q23.1	MGC4248, hypothetical protein MGC4248 MGC16186. hypothetical protein MGC16186 MGC11352, hypothetical protein MGC11352
Ctrlm57-E-5	186, 97.4 %	13q14.11	LHFP. lipoma HMGIC fusion partner
E-166m50Ctrl	181, 100 %	18q23	PTPRM, protein tyrosine phosphatase. receptor type, mu (REF?? 1 items found on Schizophrenia. 4 items found on bipolar)
Ctrlm57-E-2	163, 91 %	19q13.32	SULT2B1, sulfotransferase family, cytosolic, 2B, member
E-296_m57_Ctrl	179. 98.4 %	21q22.11	hormonally upregulated Neu-associated kinase. HUNK

Figure 2 Continued

Ctrlm57-E-4	193, 100 %	22q12.2	OSM, oncostatin M (Ref?? 2 papers found on bipolar WHAT??) LIF, leukemia inhibitory factor (cholinergic EPI64, EBP50-PDZ interactor of 64 kD SF3A1, splicing factor 3a, subunit 1, 120kD
Ctrlm57-6-E-1	155, 87.5 %	22q13.2	E1A binding protein p300, EP300
E-267_m50_Ctrl E-261_m50_Ctrl E-167m50Ctrl E-275m50Ctrl E-281m50Ctrl RevE-270m50Ctrl	191, 100 %	Yq12, Yq11.23, Yq11 223	variable charge. Y chromosome. 2 protein, VCY2
BDd_M34-14BD	187; 99%	2p23.2	BRE, brain and reproductive organ-expressed (TNFRSF1A LRRFIP1, leucine rich repeat (in FLII) interacting
BD43-10	192: 99 1% 181; 97 4%	3p22.2 3p22.1	SEC22C, vesicle trafficking protein. isoform a
E-74m43BD	195, 99.5 %	9q22 2	SHC3, neuronal Shc
BDc_M34-4BD BDc_M34-3BD	191, 100 % or 191, 100 %	11q11 11q13.4	FOLR1, folate receptor 1 precursor SKD3, suppressor of potassium transport defect 3 INPPL1, inositol polyphosphate phosphatase-like 1 FOLR2, folate receptor 2 precursor . ARIX, aristaless (Drosophila) homeobox
BD43-8	178, 100 %	11q22.3	nuclear protein, ataxia-telangiectasia locus, NPAT {Lange, 1989 #114;Weeks, 1989 #115}
E-72m43BD	160, 100 %	16q13	CNGB1, cyclic nucleotide gated channel beta 1
BD43-14	191, 100 %	16q24.2	hypothetical protein FLJ23497
E-71m39MD	147, 92 %	15q26.1	PRC1, protein regulator of cytokinesis 1
BDd_M43-19BD.	201, 100 %	19p13 11	KCNN1, potassium intermediate/small conductance (REF ?? 1 items foundon Schizophrenia. 2 items found on bipolar.

Figure 2 Continued

			SLC5A5. solute carrier family 5 (sodium iodide. IL12RB1, interleukin 12 receptor. beta 1 (41 papers found. on interleukin receptor & schizophrenia; 5 items found. on interleukin receptor & bipolar.
BDC-_M34-10BD BDC-_M34-1BD BD34-5 BD34-8 BD43-1 BD43-2 MD39-4 MD39-6 MD39-8 MD39-10 MDC-_M39-2 MDD-_M39-14 (190, 100) E-66m39MD	191, 100 %	Yq12, Yq11.23, Yq11.223	variable charge, Y chromosome, 2 protein, VCY2

Figure 2 Continued

SEQUENCE LISTING

ATTACHMENT 3. Cloned *Alu* sequences

SZ- from individuals affected with schizophrenia
CNTR- from control samples
BD - from individuals affected with bipolar disorder
MD - from individuals affected with major depression

>E-130_m37_SZ
CTGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTTGCAGACGCGTTACGT
ATCGGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTCAGCTCACCGAAACCTCCGCCTCACAGGTTCAAG
TGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAGCATTTGCCATGATACCTGGCTAATTTTGTATTTTT
AGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGAACTCCTGACCTCAGATGATCCATCTGATTTGGCC
TCCCAAACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGACCACA
CGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACTGGCCGTCGT
TTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCCCAGCT
GGCGTAATAGACGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAAGCCTG

>E-140_m48_SZ
CTATCCCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCG
TTACGTATCGGATCCAGAATTCGTGATTGCCTGTACTCCCAGCAGTTTGGGAGGCTGAGGTAGGTGGATCACGAG
GTCAGGAGTTCTAGATCAGCCTGGCCAACAGGGTGAAACCATGTCTCTACTAAAAATACAAAAATTAGTCAGGCG
TGGTGGTGGGCACCTGTAATCCCAGTTACTTGGGAGGCTGAGGCAGGAGAATTTCTTGAACCTGGAAGGCAGAGG
TTGCAGTCAGCCGAGATTGTGCAAACACCCTCCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCT
AGACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACT
GGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTT
TCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCACAGTTGCGCAGCCTGAATGGCGAATG
GAAATTGTAA

>E-150_m48_SZ
CTATGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGC
GTTACGTATCGGATCCAGAATTCGTGATTGCCTGTACTCCCAGCAGTTTGGGAGGCCAAATCAGATGGATCATCTG

Figure 3

AGGTCAGGAGTTCAAGAACCACCTTATCAACATGAAGAATCCTGGTCTCTACTAAAAGTACAAAATTAGCCAGGT
ATCATGGCAAATGCTTGTCATCCTAGCTACTCAGAAGGCTGAGGCAGAGGAATCACTTGAACCTGTGAGGCGGAG
GTTTCGGTGAGCTGAGATTGTGCAAACACCCTCCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTC
TAGACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACT
GGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCT
TTCGCCAGCTGGCGTAATAGCGAAGAGGGCCGCACCGATCGCCCTTCCAACAGTTGCGCAGCCTGAATGGCGA

> E-154_m56_SZ
ATGATTACGCCAAGCTCTAATACAACTCACTATGGGCAAATGGTCGCAACCTCGCATGCTGCATACGCGTTACGTA
TCGGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTCAGCTCACTGCAACCTCCACCTCCCAGGCTCAATG
ATCCTCCCACCTCAACTCCCCCGAGTAACTGGGACCACAGGTGCATGCCAGCATGCCCAGCTAATTTTTGTATTTT
CTGTTGAGATGGGGTTTTGCCATGTTGCCCAGGCAGGTCTCGAACTGCTGGGCTCAAGTGATCCTCCTGCCTCCAC
CTCACAAACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGACCAC
ACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACTGGCCGTCG
TTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAG
CTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGAAATT
GTAAGCGTTAATAT

> E-178_m74_SZ
AAGATCCATATGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTG
CAGACGCGTTACGTATCGGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTTGGCTCACTGCAACCTCCGC
CTCCCGGGTTCAAGAGATTCTCCTGCCTCAGCCTCCCGAGAGGCTGGGACTACAGGCATGCGCCACCATGCCCAG
CTAGTTTTTGTATTTTTAGTAGAGATGGGGTTTCCCCATGTTGGCCAGGATGATCTCGATCTCTTGACCTCGTGATC
TGCCCGCCTCAGCCTCCCAAACTTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCT
AGCTCTAGACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAA
TTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACAT
CCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAG

> E-191_m34-4_BD
ATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCGTTACGTA
TCGGATCCAGAATTCGTCGATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGACCACACGTGTGG
GGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACTGGCCGTCGTTTTACAA
CGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTA
ATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCGTGAATGGCGAATGGAAATTGTAAGCGT
TAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAA
ATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATT

Figure 3 Continued

AAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCAC
CCTAATCAAGTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGC
TTGACGGGGAAAGC

> E-221_m37_SZ
CCATATGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGAC
GCGTTACGTATCGGATCCAGAATTCGTGATTGCCTGTACTCCCAGCAGTTTGGGAGGCCAAATCAGATGGATCATC
TGAGGTCAGGAGTTCAAGAACCACCTTATCAACATGAAGAATCCTGGTCTCTACTAAAAATACAAAATTAGCCAG
GTATCATGGCAAATGCTTGTCATCCTAGCTACTCAGAAGGCTGAGGCAGAGGAATCACTTGAACCTGTGAGGCGG
AGGTTTCGGTGAGCTGAGATTGTGCAAACACCCTCCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGC
TCTAGACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCA
CTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCC
C

> E-244_m48_SZ
CCGTATGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGAC
GCGTTACGTATCGGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTCAGCTCACCGAAACCTCCGCCTCAC
AGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAGCATTTGCCATGATACCTGGCTAATT
TTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGAACTCCTGACCTCAGATGATCCATCT
GATTTGGCCTCCCAAACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCT
AGACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACT
GGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCT
TTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATG

> E-246_m48_SZ
CTATGACCATGATTACGCCAAGCTCTAATACCGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACG
CGTTACGTATCGGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTCGGCTCACTGCAACCTCCACCTCCCA
GGTTCAAGCAATTCTCCTGCCTCAGCCTCCCAAGTAGCTGAGATTACAGGCGGCTGCCATCATGCCTGGCTAATTT
TTGTATTTTTACTAAAGACGGGGTTTTGCCATGTTGGCCAGGCTGTTCTCAAACTCCTGACTTCAGGTGATCCACCT
GCCTCAGCCTCCCAAACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCT
AGACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACT
GGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCT
TTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGT

> E-251_m48_SZ

Figure 3 Continued

CATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCGTTACGT
ATCGGATCCAGAATTCGTGATTCGGAGGGTGTTTGCACAATCTTGACTAACTGCAACATCTGCCTCCCAGGTTCAA
GCAATTCTGCCTCAGCTTCCTGAGCAGCTGGGATTACAGATGAGCACTACCATGACAGGCTAATTTTTATATTTTT
AGTAGAGGCGGGGTTTCACCATGTTGGCCAGGCTGGTCATGAACTCCTGACCTCAGGTGATTCACCTGCCTCAGCC
TCCCAAACTGCTGGGAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGACCACACGTGTGGGG
GCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACTGGCCGTCGTTTTACAACG
TCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAAT
AGCGAAGAGGCCCGCACCGATCGCCCTTCC

> E-252_m48_SZ
CGATATGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGAC
GCGTTACGTATCGGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTCAGCTCACCGAAACCTCCGCCTCAC
AGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAGCATTTGCCATGATACCTGGCTAATT
TTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGAACTCCTGACCTCAGATGATCCATCT
GATTTGGCCTCCCAAACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCT
AGACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACT
GGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCC
TTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCC

> E-2531_m48_SZ
CAGCTATGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGA
CGCGTTACGTATCGGATCCAGAATTCGTGATTGCCTGTACTCCCAGCAGTTTGGGAGGCTGAGGCAGGTGAATCA
CCTGAGGTCAGGAGTTCATGACCAGCCTGGCCAACATGGTGAAACCCCGCCTCTACTAAAAATATAAAAATTAGC
CTGTCATGGTAGTGCTCATCTGTAATCCCAGCTGCTCAGGAAGCTGAGGCAGAATTGCTTGAACCTGGGAGGCAG
ATGTTGCAGTTAGTCAAGATTGTGCAAACACCCTCCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGC
TCTAGACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCA
CTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCC
CTTTCGCCAGCTGGCGTAATAGCGAAGAGGGCCGCACCGATCG

> E-2532_m48_SZ
CTATGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGC
GTTACGTATCGGATCCAGAATTCGTGATTGCCTGTACTCCCAGCAGTTTGGGAGGCTGAGGCAGGTGAATCACCTG
AGGTCAGGAGTTCATGACCAGCCTGGCCAACATGGTGAAACCCCGCCTCTACTAAAAATATAAAAATTAGCCTGT
CATGGTAGTGCTCATCTGTAATCCCAGCTGCTCAGGAAGCTGAGGCAGAATTGCTTGAACCTTGGGAGGCAGATG
TTGCAGTTAGTCAAGATTGTGCAAACACCCTCCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCT
AGACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACT

Figure 3 Continued

GGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCT
TTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTNCCAACAGTTGCGCAGCCTGAATGG

[> E-258_m48_SZ
CCATATGATCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACCGCATGCTGCAGA
CGCGTTACGTATCGGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTTGGCTCACTGCAACCTCTGCCCCC
CAGGTTCAAACGATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTATAGGCACCTGCCACCACGCCCAGCTAAT
TTTTTGCATTTTTAGTAGAGACGGGGTTTCACTATGTTGGCCAGGCTGGTCTAGAACTCCTGACCTTGTGATCCGCC
CGCCTTGGCCTCCCAAACTGCTGGGAGTAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGACC
ACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACTGGCCGT
CGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAAC

> E-261_m50_Ctrl
TGACCTTGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCGTT
ACGTATCGGATCCAGAATTCGTGATTGGAGGGTGTTTCGCACAATCTCAGCTCACCGAAACCTCCGCCTCACAGGT
TCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAGCATTTGCCATGATACCTGGCTAATTTTGT
ATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGAACTCCTGACCTCAGATGATCCATCTGATT
TGGCCTCCCAAACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGA
CCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACTGGCC
GTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCC

> E-267_m50_Ctrl
TTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCGTTACGTATCGG
ATCCAGAATTCGTGATTGCCTGTACTCCCAGCAGTTTGGGAGGCCAAATCAGATGGATCATCTGAGGTCAGGAGT
TCAAGAACCACCTTATCAACATGAAGAATCCTGGTCTCTACTAAAAATACAAAATTAGCCAGGTATCATGGCAAA
TGCTTGTCATCCTAGCTACTCAGAAGGCTGAGGCAGAGGAATCACTTGAACCTGTGAGGCGGAGGTTTCGGTGAG
CTGAGATTGTGCAAACACCCTCCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGACCACACG
TGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACTGGCCGTCGTTT
TACAACGTCGTGACTGGGAAAACCCTGGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTT

> E-269_m50_Ctrl
CTTCCAAAGGNTAAGNTCTAATATTACTCACTATAGGGAAAGCTCGGCCCCACTCATGCTGCAGACGCGTTACGT
ATTGGATCCAGAATTCGCGATTGGAGGGTGTTTGTACAATCTCTGCTCACCGAAACCTCCGCCTCACAGGTTCAAG
TGATCCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAGCATTTGCCATGATACCTGGCTAATTTTGTATTTTT
AGTAGAGACCAGGATTCTTTTATGTTGATAAGGCGGTTCTTGAACTCCTGACCTCAGATTGATTCATCTGATTTGG
CCTCCCAAACTGCTGGGAGTACAGGCAATCTGAATTCGTCAACAAGCTTCT

Figure 3 Continued

[> E-285_m56_SZ
GGTGAGAGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCGTT
ACGTATCGGATCCAGAATTCGTGATTGCCTGTACTCCCAGCAGTTTGGGAGGCTGAAGTGGGTTGATTACCCGAG
GTCAGGAGTTCCAGACCAGGTTGACCAACATGGAGAAACCCTGTCTCTACTAAAAATACAAAATTAGCCAGGTGT
ATTGGTGCGTGCCTGTATTCCCAGCTACTTGGGAGGCCGAGGCAGGAGAATCGCTGGAACCCAGGAGGCGGAGGT
TGTGGTGAGCTGAGATTGTGCAAACACCCCCCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTA
GACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACTG
GCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTT
TCGCCAGCTGGCGTAATAAGCGAAGAGGCCCGCACCGATCGCCCTTTCCAACAGTTGCGCAAGCCTGAATGGCG

> E-286_m56_SZ
GTTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCGTTACGTATCGGATCCAGAAT
TCGTGATTGGAGGGTGTTTGCACAATCTCAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCT
CAGCCTTCTGAGTAGCTAGGATGACAAGCATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGG
ATTCTTCATGTTGATAAGGTGGTTCTTGAACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCCAAACTGCTGG
GAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGACCACACGTGTGGGGGCCCG
AGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCCGCACAATTCACTGGCCGTCGTTTTACAACGTCGT
GACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCG
AAGAAGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGAAATTGTAAGCGT

> E-287_m56_SZ
TAATTAACTCACTATAGGGAAAGCTCGGGAGCACGCATGCTGCAUACGCGTTTCGTATCTGGATCCAGAATTCGC
GATTGCCTGTACTCCCAGCAGTTTGGGAGGCCAAATCAGATGGATCATCTGAGGCCAGGAGTTCAAGAACCACCT
TATCAACATGAATAATCCTGGTCTCTACTAAAAATACGAAATTAGCCAGGTATCATGGAAAATGCTTGTCATCCTA
GCTACTCAGAAGGCTGAGGCAGAGGAATCACTTGAACCTGTGAGGCGGAGGTTTCGGTGAGCTGAGATTGGGCA
AACACCCTCCAATCTGAATTCGTCCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGACCACACGCGTGGGGGCCC
GAGCTCGCGGCCGCTGTATTCTATT

> E-288_m56_SZ
GTTCAGATCTAATANGACTCACTATCGGGAAAGCTCGGCACCACGCATGCTGCAGACGCGTTACGTATCCGGATC
CATGAATTCGTGATTGCCTGTACTCCCAGCAGTTTGGGAGGCCAAATCAGATGGATCATCTGAGGTCAGGAGTTC
AAGAACCACCTTATCAACATGAAGAATCCTGGTCTCTACTAAAAATACAAAATTAGCCAGGTATCATGGCAAATG
CTTGTCATCCTAGCTACTCAGAAGGCTGAGGCAGAGGAATCACTTGAACCTGTGAGGCGGAGGTTTCGGTGAGCT
GAGATTGTGCAAACACCCTCCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGACCACACGTG

Figure 3 Continued

TGGGGGCCCGAGCTCGCGGCCGCTGCATTCTATAGTGTCACCTAAATGGCCGCACAATTCACTGGCCGTCGTTTTT
A

> E-289_m56_SZ
TTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCGTTACGTATCGG
ATCCAGAATTCGTGATTGCCTGTACTCCCAGCAGTTTGGGAGGCCAAATCAGATGGATCATCTGAGGTCAGGAGT
TCAAGAACCACCTTATCAACATGAAGAATCCTGGTCTCTACTAAAAATACAAAATTAGCCAGGTATCATGGCAAA
TGCTTGTCATCCTAGCTACTCAGAAGGCTGAGGCAGAGGAATCACTTGAACCTGTGAGGCGGAGGTTTCGGTGAG
CTGAGATTGTGCAAACACCCTCCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGACCACACG
TGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACTGGGCCGTCGTT
TTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCT
GGCGTAATAGCGAAGAGGCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGAAATTG

> E-290_m56_SZ
ATATTGATCATGATTACGCCAACGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGAC
GCGTTACGTATCGGATCCAGAATTCGTGATTGCCTGTACTCCCAGCAGTTTGGGAGGCTGAAGTGGGTTGATTACC
CGAGGTCAGGAGTTACAGACCAGGTTGACCAACATGGAGAAACCCTGTCTCTACTAAAAATACAAAATTAGCCAG
GTGTATTGGTGCGTGCCTGTAATCCCAGCTACTTGGGAGGCCGAGGCAGGAGAATCGCTGGAACCCAGGAGGCGG
AGGTTGTGGTGAGCTGAGATTGTGCAAACACCCTCCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGC
TCTAGACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCA
CTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCC
CTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCC

> E-291_m56_SZ
TGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCGTT
ACGTATCGGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTCAGCTCACCGAAACCTCCCCCTCACAGGTT
CAAGTGATTCCTCTGCCTCAGCCTTCAGAGTAGCTAGGATGACAAGCATTTGCCATGATACCTGGCTAATTTTGTA
TTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTCCTTGAACTCCTGACCTCAGATGATCCATCTGATTT
GGCCTCCCAAACTGCTGGGAGTACAGGCAATCTGAATTCCTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGAC
CACACCGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACTGGCC
GTCGTTTTACAACNTCGTGACTGGGAAAACCCTGNCGTTACCCCACTTAATCNCCCTTGCAGCACATCCCCCTTTC
GCCCAGNCTGGGCGTAATNANCGAANAGGCCCGCACCCGATCGCCC CT

> E-292_m56_SZ
ACGTCACGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCGTTACGTATCGGAT
CCAGAATTCGTGATTGCCTGTACTCCCAGCAGTTTGGGAGGGCAAATCAGATGGATCATCTGAGGTCAGGAGTTC

Figure 3 Continued

AAGAACCACCTTATCAACATGAAGAATCCTGGTCTCTACTAAAAATACAAAATTAGCCAGGTATCATGGCAAATG
CTTGTCATCCTAGCTACTCAGAAGGCTGAGGCAGAGGAATCACTTGAACCTGTGAGGCGGAGGTTTCGGTGAGCT
GAGATTGTGCAAACACCCTCCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGACCACACGTT
GTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACTGGCCGTCGTTTT
ACAACGTTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCCAG
CT

> E-293_m56_SZ
TATGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTTGCAGACGC
GTTACGTATCGGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTCAGCTCACCGAAACCTCCGCCTCACAG
GTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAGCATTTGCCATGATACCTGGCTAATTTT
GTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGAACTCCTGACCTCAGATGATCCATCTG
ATTTGGCCTCCCAAACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTA
GACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACTG
GCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCT
TTCGCCAGCTGGCGTAATAGCGAAGAGGCCGCACCCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATG

> E-294_m740_SZ
TTACGCCACGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCGTTACGTATCGG
ATCCAGAATTCGTCGATTGGAGGGTGTTTGCACAATCTCAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGAT
TCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAGCATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTA
GAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGAACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC
AAACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGACCACACGTG
TGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACTGGCCGTCGTTTTA
CAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGC
GTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAG

> E-295_m740_SZ
TATGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTTGCAGACGC
GTTACGTATCGGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTCAGCTCACCGAAACCTCCGCCTCACAG
GTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAGCATTTGCCATGATACCTGGCTAATTTT
GTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGAACTCCTGACCTCAGATGATCCATCTG
ATTTGGCCTCCCAAACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTA
GACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACT
GGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCC
TTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTTGCGCAGCCTGAA

Figure 3 Continued

> E-296_m57_Ctrl
CAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCGTTACGTATCGGATCCAG
AATTCGTGATTGGAGGGTGTTTGCACAATCTCAGCTCACTGCAACCTCTGCCTCCTGGGTTCAATTCATTCTCCTGC
CTCAGCCTTCCGAGTAGCTGGGATTACAGGCATGCCCGGCTAATTTTTGTATTTTTAGCAGAGATCGGGGTTTTGC
CATGTTGCCCAGGCTGGTCTCGAACTCCTAACCTTGTGATCTGCCCACCTCGGCCTCCCAAACTGCTGGGAGTACA
GGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGACCACACGTGTGGGGGCCCGAGCTCGC
GGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACTGGCCCGTCGTTTTACAACGTCGTGACTG
GGAAAACCCTGGGCGTTACCCAACTTAATCGCCCTTGCAGCACATCCCCCTTTCGCCAGCTTGGC

> E-297_m740_SZ
TATGACCATCATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCG
TTACGTATCGGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTCAGCTCACCGAAACCTCCGCCTCACAGG
TTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAGCATTTGCCATGATACCTGGCTAATTTTGT
ATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGAACTCCTGACCTCAGATGATCCATCTGATT
TGGCCTCCCAAACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGA
CCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACTGGCC
GTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCG
CCAGCTGGCGTAATAGCGAAGAGGCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGAA
ATT

> E-298_m57_Ctrl
GTCCCGATCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCGTTACGTATCGGATCC
AGAATTCGTGATTGGAGGGTGTTTGCACAATCTCAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTC
TGCCTCAGCCTTCTGAGTAGCTAGGATGACAAGCATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGA
CCAGGATTCTTCATCGTTGATAAGGTGGTTCTTGAACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCCAAA
CTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGACCACACGTGTGG
GGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCCTAAATGGCCGCACAATTCACTGGCCGTCGTTTTAC
AACGTCGTGACTGGGAAAACCCTGGGCGTTACCCCAACTTAATCG

|> E-299_m57_Ctrl
GTCAAGATCGAATAGGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCCGACGCGTTACGTATCGGATCC
AGAATTCGTGATTGCCTGTACTCCCAGCACTTTGGGAGGCAAATCAGATGGATCATCTGAGGTCAGGAGTTCAA
GAACCATCCTTATCAACATGAAGAATCCTGGTCTCTACTAAAAATACAACATTAGCCAGGTATCATGGCAAATGC
TTGTCATCCTAGCTACTCACAAGGCTGAGGCAGAGGAATCACTTGAACCTGTGAGGCGCAGGTTTCGGTGAGCTG

Figure 3 Continued

AGATTGTGCAAACACCCTCCAATCTGAATTCGTCGACAAGCTCTCTCGAGCCTAGGCTAGCTTTAGANCACACGTG
TGGGGGC

5> E-300_m57_Ctrl
GTTGAAACGGCAAGATCTAATACGACTCACTATAGGGAAAGCTCGGCACTACGCATGCTGCAGACGTGTTGACGT
ATCGGATCCAGAATTCGTGATTGGAGGGCGTTTGCGCAATCTTGACTAACTGCAACATCTGCCTCCCAGGCTCAAG
CAATTCTGCCTCAGTTTTCTGAGCAGCTGGGATTACAGATGAGCACTACCATGACAGGCTAATTTTTATATTTTTAC
TAGAGGCGGGGATTCACCATGTCGGCCAGGTTGGTCATGAACTCCTGACCTCAGGCGATTCACCTGCCTCCGCCTC
CCAAACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAG

> E-304_m57_Ctrl
CTACGTACGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCGTTACGTATCGGA
TCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTCAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTC
CTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAGCATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGA
GACCAGGATTCTTCATGTTGATAAGGCGGTTCTTGAACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCCAA
ACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGACCACACGTGTG
GGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCCGCACAATTCACTGGCCGTCGTTTTAC
AACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGC

> E-305_m740_SZ
TTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCGTTACGTATCGG
ATCCAGAATTCGCGATTGGAGGGTGTTTGCACAATCTCAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATT
CCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAGCATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAG
AGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGAACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCCA
AACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCCGAGCCTAGGCTAGCTCTAGACCACACGTG
TGGGGGCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACTGGCCGTCGTTTTTA
CAACG

> E-308_m74_SZ
TTACGTCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCGTTACGTATCGG
ATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTCAGCTCACCGAAATCTCCGCCTCACAGGTTCAAGTGATT
CCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAGCATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAG
AGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGAACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCCA
AACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGACCACACGTGT
GGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACTGGCCGTCGTTT

Figure 3 Continued

|> E-309_m74_SZ
AGGCAAGATCTAATACGACTCACTATAGGGAAACGCTCGGTACCACGCATGCTGCAGACGCGTTACGTATCGGAT
CCAGAATTCGTGATTGCCTGTACTCCCACGCAGTTTGGGAGGCCAAATCAGATGGATCATCTGAGGTCAGGAGTT
CAAGAACCACCTTATCAACATGAAGAATCCTGGTCTCTACTAAAAATACAACATTAGCCAGGTATCATGGCAAAT
GCTTGTCATCCTAGCTACTCAGAAGGCTGAGGCAGAGGAATCACTTGAACCTGTGAGGCGGAGGTTTCGGTGAGC
TGAGATTGCGCAAACACCCTCCAATCTGAATTCCTCTGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGACCCCACG
TGTGGGGGCCCGAGCTCGCCGTCGCTGTATTTCTATAGTCGTC

> E-310_m74_SZ
TTACGTCACCGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCGTTACGTATCG
GATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTCAGCTCACTGCAACCTCTGCCTCTCAGGTTCAAGTGAT
TCTCCTGCCTCATCCTCCCCAGTAGCTGGGTTTACAGGCATGCACCACCACAGCTGGCTAATTTTTTGTATTTTTAGT
AGAGATGGGGTTTCACCATGTTGGACAGGCTAGTCTTGAACTCCTGACCTCAAGTGATCCACCCGCCTCAGCCTCT
CAAACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGACCACACGT
GTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGGCCGCACAATTCACTGGCCGTCCGTT
TTACAACGTCCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAA

> E-311_m74_SZ
AAACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCGTTACGTATCGG
ATCCAGAATTCGTGATTGCCTGTACTCCCAGCAGTTTGGGAGGCCGAGGTGGGTGGATCACCTGAGGCTGAGAGT
TCGAGACCAGCCTAGCCAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAAATTAGCCAGGCAAGGCAGCAC
ACGCCTGTAATTCCACCTACTCGGGATGCTGAGGCATGAGAATCGCTTGAACCTGGGAGGTGGAGCTTGCAGTGA
ACTGAGATTGTGCAAACACCCTCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGACCACACG
TGTGGGGGCCCGAGCTCGCCGGCCGCTGTATTCTATTAGTGTCACCTAAATGGGCCGCACAATTCACTGGCCGTCC
GTTTTACAACGTCGTGACTGGGAAAACC

> E-312_m74_SZ
CGAATACGACTACTATACGGAAAGCTCGGTACCACGCATGCTGCACACGCGTTACGCATCGGATCCAGAATTCGT
GATTGCCTGTACTCCCAGCAGTTTGGGAGGGCAAATCAGATGGATCATCTGAGGTCAGGAGTTCAAGAACCACCT
TATCAACATGAAGAATCCTGGTCTCTACTAAAAATACAAAATTAGCCAGGTATCATCGGCAAATGCTTCGTCATCC
TAGCTACTCAGAAGGCTGAGGCAGAGGAGTCACTTGAACCTGTGAGGCGGAGGAAACGGCGAGATGAGATTGTG
CAAACACCCTCCAATTTGAAATTCGTCGACAAGCTTCTCCGAGCTCTAGGCTAGCTCTAGACCCACACGTGTGGGG
GCCCCGAGCTCGCG

> E-313_m74_SZ

Figure 3 Continued

TATGACATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCGT
TACGTATCGGATCCAGAATTCGTGATTGCCTGTACTCCCAGCAGTTTGGGAGGCTGAGACAGGTGGAACACTTGA
GGCCAGGAGTTTGCAACCAGCCTGGCCAACATGGTGAAACCCTATCTCTACCACAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAATTAGCCTGGCATGGTGGTGCGTGCCTGTAATCCCAGCTACTCAGGAGGCTGAGGCACGAGAATC
GCTTGAACCCGGTGGGCAAGGGTTGCAGCGATCCGAGATTGTGCAAACACCCTCCAATCTGAATTCGTCGACAAG
CTTCTCGAGCCTAGGCTAGCTCTAGACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCAC
CTAAATGGCCGCACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAA
TCGCCTTGCAGCACATCCCCC

> E-315_m74_SZ
TGaTTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCGTTACGTATC
GGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTCGGCTCACTGCAACTTCTGCCTCCTGGGTTCACACTG
TTCTCCTGCCTAAGCCTCCCAAGTAGCTGGGACTACAGGCGCGTGCCACCATGCCCGGCTAATTTTTTGTATTTTTA
GTAGAGAAGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGATCTCCTGATATTGTGATCCACCCGCCTCGGCCTCT
CAAACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGACCACACGT
GTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACTGGCCGTCGTTTT
ACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGG
CGTAATAGCGAAGAGGCCGCACCGATCGCCCTTCCCAACAGTTGCGC

> E-314_m74_SZ
ATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCGTTACGTATCG
GATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTCGGCTCACTGCAACTTCTGCCTCCTGGGTTCACACTGT
TCTCCTGCCTAAGCCTCCCAAGTAGCTGGGACTACAGGCGCGTGCCACCATGCCCGGCTAATTTTTTGTATTTTTA
GTAGAGAAGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGATCTCCTGATATTGTGATCCACCCGCCTCGGCCTCT
CAAACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGACCACACGT
GTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACTGGCCGTCGTTTT
ACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGG
CGTAATAGCGAAGAGGCCGCACCGATCGCCCTT

4> E-319_m74_SZ
TATGACCATGATTACGCCAAGCTCTAATACCGACTCACTATAGGGAAACGCTCGGTACCACGCATGCTGCAGACG
CGTTACGTATCGGATCCAGAATTCGTGATTGCCTGTACTCCCAGCAGTTTGGGAGGCCGAGGTGGGTGGATCACCT
GAGGTCAGGAGTTCGAGACCAGCCTGGCCAACGTAGTGAAAACCCCATCTCTACTAAAAATACAAAAAAACTTAG
CCAGGGGTGGTGGTGGGCACCTATAATCCCAGCTACTTAGGAGGCTGAGGCTGGAGAATCGTTTGAACCTGGGAG
GGAGAGGTTGCAGTGAGCTGAGATTGTGCAAACACCCTCCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGG
CTAGCTCTAGACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCAC

Figure 3 Continued

AATTCACTGGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCA
CATCCCCCTTTCGCCAGCTGGCGTAATAACGAAGAGGCCGCACCGA

9[]> E-320_m74_SZ
ATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCGTTACGTA
TCGGATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGACCACACGTGTGGGGGCCCGAGCTCGC
GGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACTGGCCGTCGTTTTACAACGTCG
TGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGC
GAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGAAATTGTAAGCGTTAATA
TTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCT
TATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAAC
GTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATCA
AGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCC

8> E-321_m74_SZ
TATGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCG
TTACGTATCGGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTCGGCTCACTGCAACTTCTGCCTCCTGGG
TTCACACTGTTCTCCTGCCTAAGCCTCCCAAGTAGCTGGGACTACAGGCGCGTGCCACCATGCCCGGCTAATTTTT
TGTATTTTTAGTAGAGAAGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGATCTCCTGATATTGTGATCCACCCGC
CTCGGCCTCTCAAACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAG
ACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACTGG
GCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTT
TCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATG

3> E-322_m74_SZ
ACGTACGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCGTTACGTATCGGATC
CAGAATTCGTGATTGGAGGGTGTTTGCACAATCTTGGCTCACTGTAACCTCTGCCTCCTGGGTTCAAGTAATTCTC
CTGTCTCAGCCTCCTGAGTAGCTAGGATTACTGGTGCCCGCCACCATGCCCGGCGAATTTTTGTATTTTTAGTAGA
GATGGGGTTTCACTATGTTGCCCAGGGTGGTCTCAAACTCCTGACCTCAAGTGATCCACCTGCTTCAGCTTCCCAA
ACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGACCACACGTGTG
GGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACTGGCCGTCGTTTTACA
ACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCTTGCAGCACATCCCCCCTTTCGCCAGCTGGCGT
AATAGCGAAGAGGCCCGCACCCGATCGCCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGAAATT

7[]> E-323_m74_SZ

Figure 3 Continued

AAACGCAAGCTCTAATACGACTCACTATAGGGAAAGTTCGGTACCACGCATGCTGCAGACGCGTTACGTATCGGA
TCCAGAATTCGTGATTGCCTGTACTCCCAGCACGTTTGGGAAGCCGAGGTGGGAAGATCGCTTCGAGGTCAGGAG
TTCAAGACCAGCCTGGCCAACATGGCAAAACCTCGTCTCTACTAAAAATACAAAACTTAGCCAGGCCGTGTTGGC
ATCGCACCCATAGTCCCTGCTAATCAGGAGGCTGAGGCTTGAACATGGGAGGTGGAGGCTGCAGTGAGCTGAGAT
TGTGCAAACACCCTCCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGACCACACGTGTGGGG
GCCCGAGCTCGCGGCCGCTGTATTCTATAGT

4> E-324_m74_SZ
GTTAAGATCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCGTTACGTATCGGATCC
AGAATTCGTGATTGGAGGGTGTTTGCACAATCTCAGCTCACTGCAACCTCCACCTCTACGACTCAAGTGATTATCC
CACCTCAACCTCCCAAGTAGCAGGGACTGAAGGTGTGCTTTGCCACGCCCAGCTAATTTTTTGTATTTTTTGTAGA
GACGGATTTTCACCATGTAGCCCAGGCTGGTCTCAAACTCCTGAGCTTAAGCGATCCACCTTCCTGGACCTCCCAA
ACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGACCACACGTGTG
GGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGGCCGCACAATTCACTGGCCGTCGTTTTAC
AACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCG
TAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCACAGTTGCGCAGCCTGAATGGCGAATGGAAATTTAA

1]> E-325_m74_SZ
CAGCTATGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGA
CGCGTTACGTATCGGATCCAGAATTCGTGATTTGCCTTGTACTCCCAGCAGTTTGGGAGGCTGAGGCAGGTGAATC
ACCTGAGGTCAGGAGTTCATGACCAGCCTGGCCAACATGGTGAAACCCCGCCTCTACTAAAAATATAAAAATTAG
CCTGTCATGGTAGTGCTCATCTGTAATCCCAGCTGCTCAGGAAGCTGAGGCAGAATTTGCTTGAACCTGGGAGGC
AGATGTTGCAGTTAGTCAAGATTGTGCAAACACCCTCCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTA
GCTCTAGACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATT
CACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCTGGCGTTAC

3]> E-149_m48_SZ
ACGCTTCCAAGGATTCAACAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGC
GTTACGTATCGGATCCAGAATTCGTGATTAGGGTGTTTGCACAATCTCGGCTCATTGTAACCTCTGCCTCCCAGGT
TGCAGTGATTCTCCTGTCTCAGCCTCCCAAGTAGCTGGCATTACAGGTTCCCACCACTACACCCAACTAATTTTTGT
ATTTTTAGCAGAAATGGGGTTTCCCCATGTTGACCTGGCTGGTCTCGAACTCCTGACCTTGTGATCTGCCCGCCTTG
GCCTCCCAAACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGACC
ACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACTGGCCGT
CGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCC
AGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGA

Figure 3 Continued

3> E-302_m57_Ctrl
GATTACGCCAAGCTCTAATACTACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCGTTACGTATC
GGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTCAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTG
ATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGACGACAAGCATTTGCCATGATACCTGGCTAATTTTGTATTTTTAG
TAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGAACTCCTGACCTCAGATGATCCACCTGATTTGGCCTC
CCAAACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGACCACACG
TGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACTGGCCGTCGTTT
TACAACGTCGTGACTGGGAAAACCTG

5> E-119m57Ctrl
CAGCTATGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGA
CGCGTTACGTATCGGATCCAGAATTCGTGATTGCCTGTACTCCCAGCAGTTTGGGAGGCAGAGGCAGGTGGATCA
CCTGAGGTCGGGAGTTCGAGAACCGCCTGACCAACATGGAGAAACCCCGTCTCTGCTAAAAATACAAAATTAGCT
AGGTATGGTGGTACTTGCCCGTAATCCCAGCTATTCAGAAGGCTGAGGCAGGAGAGTCACTTGAACCCAGGAGTC
AGAGGTTGCAGTCAGCTGAGATTGTGCAAACACCCTCCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCT
AGCTCTAGACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAA
TTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACAT
CCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAAT
GGCGAATGG

[|> E-120m57Ctrl
AATAGCTATGCCCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGTATGCTCGGAGCTAGGCATGCTGCA
GACGCGTTACGCATTACGATCCAGAATCCAGAGATTGGAGGTGGCTGGCGTAATATCGGTTTAGTGGGACCTGTG
CCTCCGGGTTCCAGGTGTTGCTAGTGTTTGAACCTCCTGAGCATCATTGGATAACAGTAGCCTCTCACCATGCTCA
TCTTGTGCTTGTATTGGTGGCAGCGGTCCACCATGCCGGTTATGCTGAACTCGGACTCATCACCTTAAATTAACCA
CCTGCCTCAGACTCCGAAACTGCTGGTAGTACAGGCAATCTGCATTCGTCTGCATTCTTCTACAGCCTAGGCTAGC
TATAGACCACACTTGACCACGGCCCGAGCTCCCGGCCGCTTGGATTCTATAGTGTCATATAAAGGCCCGAACAATT
CACTGCACCGTAGTTT
Sorry. no matches found

5> E-166m50Ctrl
AACAGCTATGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCA
GACGCGTTACGTATCGGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTCGGCCCACTGCAACCTCCGCCT
CCCGGGTGCAAGCAGTTCTCCTACCTCAGCCTCCTGAGTAGCTAGGATTACAGGCACACCTGGCTAATTTTGTGGT
TTTAGTAGAGACGGCGTTTCACCATGTTGGCTAGGCTGGTCTCGAACTCCTCACCTCAAATGATCCACCTGCCTCA

Figure 3 Continued

60 6818.060602

6> E-271m50Ctrl
TTGCCCATGCTTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCGTT
ACGTATCGGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTCAGCTCACCATGACCTCTGCCTCCTGGGTT
CAAGCGATTCTCTGGACTCAGCCTCCTGAGTAGCTGGAATTACAGGGATTCGCCACCATGCCCAGCTAATTTTGTA
TGTTTAGTAGAGACAGGGTTTCTCCAAATTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGGTGATCCGCCCGCCT
TGGCCTCCCAAACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGA
CCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACTGGCC
GTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCG
CCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGA

0> E-272m50Ctrl
CAATACGCTTGACCATGATTACGCCAAGCTCTAATACGACTACTATAGGGAAAGCTCGGTACCACGCATGCTGC
AGACGCGTTACGTATCGGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTCAGCTCACTGCAGCCTCCTCC
CTCTGAGGTCAAGTGATACTGCTGCCTCAGCCTCCTGAGTAGCTGGGATTACAGGCACCCACCACCAACCCTGGCC
AATTTTTGTATTTTTAGTAGAGACAGAGTTTCACCATGCTGGCCAGGCTGGTCTCAAACTCCTGCCCTCAGATGTTC
CACCCACCTTGGCCTCCCAAACTGCTGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTA
GCTCTAGACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATT
CACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCC
CCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCAACAGTTGCGCAGCCTG

2> E-273m50Ctrl
GCTCGGTACCACGCATGCTGCAGACGCGTTACGTATCGGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCT
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTTATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCCAAACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACA
AGCTTCTCGAGCCTAGGCTAGCTCTAGACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTC
ACCTAAATGGCCGCACAATTCACTGGCCGGCGTTTTACAACGTCGCGACTGGGAAAACCCTGGCGTTACCCAACT
TAATCGCCTTGCAGCACATCCCC

2> E-275m50Ctrl
ACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCGTTAC
GTATCGGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTCAGCTCACCGAAACCTCCGCCTCACAGGTTCA
AGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAGCATTTGCCATGATACCTGGCTAATTTTGTATTT
TTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGAACTCCTGACCTCAGATGATCCATCTGATTTGG
CCTCCCAAACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGACCA
CACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACTGGCCGTC

Figure 3 Continued

GTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCA
GCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAAT

0> E-279m50Ctrl
AAGACCATGATAACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCGT
TACGTATCGGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTCAGCTCACTGCAGCCTCCTCCCTCTGAGG
TCAAGTGATTCTGCTGCCTCAGCCTCCTGAGTAGCTGGGATTACAGGCACCCACCACCAACCCTGGCCAATTTTTG
TATTTTTAGTAGAGACAGAGTTTCACCATGCTGGCCAGGCTGGTCTCAAACTCCTGCCCTCAGATGTTCCACCCAC
CTTGGCCTCCCAAACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAG
ACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACTGG
CCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTT
CGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATG
GAAAT

2> E-281m50Ctrl
AACAGCTATGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCA
GACGCGTTACGTATCGGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTCAGCTCACCGAAACCTCCGCCT
CACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAGCATTTGCCATGATACCTGGCTA
ATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGAACTCCTGACCTCAGATGATCCA
TCTGATTTGGCCTCCCAAACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGC
TCTAGACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCA
CTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCC
CTTTCGCCAGCTGGCGTAATAACGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAA

2> E-283m56SZ
AACAGCTATGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCA
GACGCGTTACGTATCGGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTTGGCTCACTGTAACCTCTGCCT
CTTGGGTTCAAGTAATTCTCCTGTCTCAGCCTCCTGAGTAGCTAGGATTACTGGTGCCCGCCACCATGCCCGGCAA
ATTTTTGTATTTTTAGTAGAGATGGGGTTTCACTATGTTGCCCAGGGTGGTCTCAAACTCCTGACCTCAAGTGATCC
ACCTGCTTCAGCTTCCCAAACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAG
CTCTAGACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTC
ACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCC
CCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCTTCCCAACAGTTGCGCAGCCTGA

2> E-284m56SZ

Figure 3 Continued

AGCTATGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGAC
GCGTTACGTATCGGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTCAGCTCACCGAAACCTCCGCCTCAC
AGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAGCATTTGCCATGATACCTGGCTAATT
TTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGAACTCCTGACCTCAGATGATCCATCT
GATTTGGCCTCCCAAACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCT
AGACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATTGGCCGCACAATTCACT
GGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCT
TTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAAT
GGAAATTGTAAGCG

7> E-61m34BD
TTAAACAGCTATGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCT
GCAGACGCGTTACGTATCGGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTCGGTTCACTGCAACTTCTG
CCTCCCAGGTTCAAGCAATTATCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGTGCCCGCCACCACACTCAGC
TAATTTTCGTATTTTTAGTAGAGACGGTTTCACCATCTTGGCTAGGCTGGTCTTGAGCTCCTGACTGCGTGATCCAC
CCGCCTTGGCCCCCCAAACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCT
CTAGACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCA
CTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCC
CTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGA

2> E-62m34BD
CTTGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCG
TTACGTATCGGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTTGGCTCACTGTAACCTCTGCCTCCTGGG
TTCAAGTAATTCTCCTGTCTCAGCCTCCTGAGTAGCTAGGATTACTGGTGCCCGCCACCATGCCCGGCAAATTTTT
GTATTTTTAGTAGAGATGGGGTTTCACTATGTTGCCCAGGGTGGTCTCAAACTCCTGACCTCAAGTGATCCACCTG
CTTCAGCTTCCCAAACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTA
GACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACTG
GCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTT
TCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAAT
GGAAATTGTAAGCGTTAATATT

2> E-63m34BD
AACAGCTATGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCA
GACGCGTTACGTATCGGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTCAGCTCACCGAAACCTCCGCCT
CACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAGCATTTGCCATGATACCTGGCTA
ATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGAACTCCTGACCTCAGATGATCCA

Figure 3 Continued

TCTGACTTGGCCTCCCAAACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGC
TCTAGACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCA
CTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCC
CTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCTTCCCAACAGTTGCGCAGCCTGAATGGCGA
ATG

2> E-66m39MD
TATGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCG
TTACGTATCGGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTCAGCTCACCGAAACCTCCGCCTCACAGG
TTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAGCATTTGCCATGATACCTGGCTAATTTTGT
ATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGAACTCCTGACCTCAGATGATCCATCTGATT
TGGCCTCCCAAACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTATAGA
CCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACTGGCC
GTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCTTGCAGCACATCCCCTTTCGCC
AGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAA

2> E-68m39MD
CAGCTATGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGA
CGCGTTACGTATCGGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTCAGCTCACCGAAACCTCCGCCTCA
CAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAGCATTTGCCATGATACCTGGCTAAT
TTTGTATTTTTAGTAGAGGCCAGGATTCTTCATGTTGATAAGGTGGTTCTTGAACTCCTGACCTCAGATGATCCATC
TGATTTGGCCTCCCAAACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTC
TAGACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACT
GGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCT
TTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAAT
GGAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAACTCATTTTTTAACCAA

3> E-71m39MD
AAGATTGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGAC
GCGTTACGTATCGGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTCAGCTCACTGCAACCTTCACCTCCC
AGGTTCAAGCGATTCTCATGCCTCAGCCTTCCGAATAGTTGAGATTACAGGCTCGTGCCACCACACCCAGCTAATT
TTTTGTATTTTTAGTAGAGATGGGGTTTCACCATGTTGGCCAGGCTGGTCTTGAGCTCCTGACCTCAAGTAATCTGC
CCACCTCAGCCTCCAAAACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGC
TCTAGACCACACGTGTGGGGGCCCGAGCTCGCGGCCGATGTATTCTATAGTGTCACCTAAATGGCCGCACAATTC
ACTGGCCGTCGTTTTACAACGTCGAGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCC
CCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGACCTTTC

Figure 3 Continued

> E-72m43BD
TAAACACGTTGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGC
AGACGCGTTACGTATCGGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTCGGCTCACTGCAACATCCGCC
TCCCGAGTAGCTGGGACCACAGGTGTGCACCACCTTTCCGGGCTAATTTTTGTATTTTTAGTAGAGACAGGGTTTT
GCCATGTTGGTCAGGCTGGTCTTGAACTCCTGACCTCAGGTGATTTGCCCACCTCAGCCTCCCAAACTGCTGGGAG
TACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGACCACACGTGTGGGGGCCCGAGC
TCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACTGGCCGTCGTTTTACAACGTCGTGACT
GGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAG
AGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGAAATTGTAAGCGTTAATA

> E-74m43BD
AAACAGCTATGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGC
AGACGCGTTACGTATCGGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTCAGCTCATTGCGAGCTCCACC
TCCCAGGTTCAAGCAATTCTCCTACCTCAGCAACTCCTGAGTAGCTGAGACTACAGGTGTGTGCCACTATGCCTGG
CTAACTTTTTTTGTATTTTTAGTAGAGACAGGGTTTCACCATGTTGGCCAGGCTAGTCTCGAACACCTGACCTCAG
ATGATCCACCTGCCTCGGCCTCCCAAACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCT
AGGCTAGCTCTAGACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCG
CACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAG
CACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCT
GAATGGCGAATGGAAAT

> E-75m43BD
CAGCTATGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGA
CGCGTTACGTATCGGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTTGGTTCACTACAACCTCCAATCTC
CAGGTTCAAGGATTCTCCTGCCTCAGACTCCTGAGTAGCTGGGATTACAGGCATCCACCAACATGCCTGGCTAATT
TTTTTATTTTTAGCAGAGACGGGGTTTTGCCATATTGGCCATGCTGGTCTCAAACTCCTGACCTCATGTGATCCACC
CGCCTTGGCCTCCCAAACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTC
TAGACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATTCACT
GGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCT
TTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAA
TGGAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCAT

> E-77m43BD
CAGCTAACAGCTATGACCTGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGC
TGCAGACGCGTTACGTATCGGATCCAGAATTCGTGATTGCCTGTACTCCCAGCAGTTTCGGAGGTTGAGGCGGGTG

Figure 3 Continued

603 6818 .060602

GATTACCTGAGGTCAGGAGTTTAAGATCAGCCTGGCCAACCTGATGAAACCCCATCTCTACTAAAAATACAAAAA
ATTAGCCTGGTGTGTTGGTGGGCATCTGTAATCCCAGCTACTCGGGAGGCTGAGGCAGGATAATCACTTGAACCT
GGGAGGTGGTGGTTGCAGTGAGCTGAGATTGTGCAAACACCCTCCAATCTGAATTCGTCGACAAGCTTCTCGAGC
CTAGGCTAGCTCTAGACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCC
GCACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCA
GCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCC

> E-78m43BD
ACAGCTATGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAG
ACGCGTTACGTATCGGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTCGGCTCAATGCAACCTCAGCCTC
CTGGGTTCAAGCAATTCTCCTGTCTCAGCCTCCCGAGTAGCTGGGATTACAGGCACATGCCACCATGCCCAACTAA
TTTTTGTATTTTTAGTAGAGACAGGGTTTTGCCATGTTGGCCAGGCTGGTCTCAAACTCCTGACCTCAGGTGGTCCA
CCGGCCTCAGCCTCCCAAACTGCTGGGAGTACAGGCCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTA
GCTCTAGACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATT
CACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCC
CCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCTTCCAACAGTTGCGCAGCCTGAA

> E-79m43BD
AACAGCTATGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCA
GACGCGTTACGTATCGGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTCAGCTCACTGCAACCTCCGTTT
CCCAGGTGCAACCGATTCTCCTGCCTCAGACCTCTGAAGCGGCTGGGACTACAGGTGCCTGCCACCTCACCCGGCT
AATTTTTGTATTTTTAGTAAGAGATGGGGTTTCACCACATTGGCCGGGGTGGTCTCAAACTCCTGACCTCAAGTGA
TCCTTCCATCTTGGCCTCCCAAACTGCTGGGAGTACAGGCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGC
TAGCTCTATACCACACGTGTGGGGGCCCGAGCTCCGCGGCCGCTGTATTCTATAGTGTTACCTAAATGGCCGGACA
ATTCACTGGCCGTCGGTTTACAACGTCAGGACTGGGAAAACCCTGGCGTTACCCAACTTAATGCC

> E-81m43BD
CAGCTATGACCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGA
CGCGTTACGTATCGGATCCAGAATTCGTGATTGGAGGGTGTTTGCACAATCTCGGCTCAATGCAACCTCAGCCTCC
TGGGTTCAAGCAATTCTCCTGTCTCAGCCTCCCGAGTAGCTGGGATTACAGGCACATGCCACCATGCCCAACTAAT
TTTTGTATTTTTAGTAGAGACAGGGTTTTGCCATGTTGGCCAGGCTGGTCTCAAACTCCTGACCTCAGGTGGTCCA
CCGGCCTCAGCCTCCCAAACTGCTGGGAGTACAGGCCAATCTGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTA
GCTCTAGACCACACGTGTGGGGGCCCGAGCTCGCGGCCGCTGTATTCTATAGTGTCACCTAAATGGCCGCACAATT
CACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCC
CCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCTTCCAACAGTTGCG

Figure 3 Continued

\> E-167m50Ctrl
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

\> E-271m50Ctrl
CAGCTCACCATGACCTCTGCCTCCTGGGTTCAAGCGATTCTCTGGACTCAGCCTCCTGAGTAGCTGGAATTACAGG
GATTCGCCACCATGCCCAGCTAATTTTGTATGTTTAGTAGAGACAGGGTTTCTCCAAATTGGTCAGGCTGGTCTCG
AACTCCCGACCTCAGGTGATCCGCCCGCCTTGGCCTCCC

\> E-272m50Ctrl
CAGCTCACTGCAGCCTCCTCCCTCTGAGGTCAAGTGATACTGCTGCCTCAGCCTCCTGAGTAGCTGGGATTACAGG
CACCCACCACCAACCCTGGCCAATTTTTGTATTTTTAGTAGAGACAGAGTTTCACCATGCTGGCCAGGCTGGTCTC
AAACTCCTGCCCTCAGATGTTCCACCCACCTTGGCCTCCC

\> E-273m50Ctrl
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTTATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

\> E-275m50Ctrl
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

\> E-279m50Ctrl
CAGCTCACTGCAGCCTCCTCCCTCTGAGGTCAAGTGATTCTGCTGCCTCAGCCTCCTGAGTAGCTGGGATTACAGG
CACCCACCACCAACCCTGGCCAATTTTTGTATTTTTAGTAGAGACAGAGTTTCACCATGCTGGCCAGGCTGGTCTC
AAACTCCTGCCCTCAGATGTTCCACCCACCTTGGCCTCCC

\> E-281m50Ctrl
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

\> E-283m56SZ

Figure 3 Continued

TGGCTCACTGTAACCTCTGCCTCTTGGGTTCAAGTAATTCTCCTGTCTCAGCCTCCTGAGTAGCTAGGATTACTGGT
GCCCCGCCACCATGCCCGGCAAATTTTTGTATTTTTAGTAGAGATGGGGTTTCACTATGTTGCCCAGGGTGGTCTCA
AACTCCTGACCTCAAGTGATCCACCTGCTTCAGCTTCCC

> E-284m56SZ
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

> E-61m34BD
CGGTTCACTGCAACTTCTGCCTCCCAGGTTCAAGCAATTATCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGT
GCCCGCCACCACACTCAGCTAATTTTCGTATTTTTAGTAGAGACGGTTTCACCATCTTGGCTAGGCTGGTCTTGAG
CTCCTGACTGCGTGATCCACCCGCCTTGGCCCCCC

2> E-62m34BD
TGGCTCACTGTAACCTCTGCCTCCTGGGTTCAAGTAATTCTCCTGTCTCAGCCTCCTGAGTAGCTAGGATTACTGGT
GCCCGCCACCATGCCCGGCAAATTTTTGTATTTTTAGTAGAGATGGGGTTTCACTATGTTGCCCAGGGTGGTCTCA
AACTCCTGACCTCAAGTGATCCACCTGCTTCAGCTTCCC

2> E-63m34BD
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGACTTGGCCTCCC

2> E-66m39MD
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

2> E-68m39MD
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGGCCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

Figure 3 Continued

3> E-71m39MD
CAGCTCACTGCAACCTTCACCTCCCAGGTTCAAGCGATTCTCATGCCTCAGCCTTCCGAATAGTTGAGATTACAGG
CTCGTGCCACCACACCCAGCTAATTTTTTGTATTTTTAGTAGAGATGGGGTTTCACCATGTTGGCCAGGCTGGTCTT
GAGCTCCTGACCTCAAGTAATCTGCCCACCTCAGCCTCCA

4> E-72m43BD
CGGCTCACTGCAACATCCGCCTCCCGAGTAGCTGGGACCACAGGTGTGCACCACCTTTCCGGGCTAATTTTTGTAT
TTTTAGTAGAGACAGGGTTTTGCCATGTTGGTCAGGCTGGTCTTGAACTCCTGACCTCAGGTGATTTGCCCACCTC
AGCCTCCC

3> E-74m43BD
CAGCTCATTGCGAGCTCCACCTCCCAGGTTCAAGCAATTCTCCTACCTCAGCAACTCCTGAGTAGCTGAGACTACA
GGTGTGTGCCACTATGCCTGGCTAACTTTTTTTGTATTTTTAGTAGAGACAGGGTTTCACCATGTTGGCCAGGCTAG
TCTCGAACACCTGACCTCAGATGATCCACCTGCCTCGGCCTCCC

2> E-75m43BD
TGGTTCACTACAACCTCCAATCTCCAGGTTCAAGGATTCTCCTGCCTCAGACTCCTGAGTAGCTGGGATTACAGGC
ATCCACCAACATGCCTGGCTAATTTTTTTATTTTTAGCAGAGACGGGGTTTTGCCATATTGGCCATGCTGGTCTCAA
ACTCCTGACCTCATGTGATCCACCCGCCTTGGCCTCCC

3> E-78m43BD
CGGCTCAATGCAACCTCAGCCTCCTGGGTTCAAGCAATTCTCCTGTCTCAGCCTCCCGAGTAGCTGGGATTACAGG
CACATGCCACCATGCCCAACTAATTTTTGTATTTTTAGTAGAGACAGGGTTTTGCCATGTTGGCCAGGCTGGTCTC
AAACTCCTGACCTCAGGTGGTCCACCGGCCTCAGCCTCCC

5> E-79m43BD
CAGCTCACTGCAACCTCCGTTTCCCAGGTGCAACCGATTCTCCTGCCTCAGACCTCTGAAGCGGCTGGGACTACAG
GTGCCTGCCACCTCACCCGGCTAATTTTTGTATTTTTAGTAAGAGATGGGGTTTCACCACATTGGCCGGGGTGGTC
TCAAACTCCTGACCTCAAGTGATCCTTCCATCTTGGCCTCCC

3> E-83m43BD
CGGCTCAATGCAACCTCAGCCTCCTGGGTTCAAGCAATTCTCCTGTCTCAGCCTCCCGAGTAGCTGGGATTACAGG
CACATGCCACCATGCCCAACTAATTTTTGTATTTTTAGTAGAGACAGGGTTTTGCCATGTTGGCCAGGCTGGTCTC
AAACTCCTGACCTCAGGTGGTCCACCGGCCTCAGCCTCCC

Figure 3 Continued

[]10> E-120m57Ctrl
AATAGCTATGCCCATGATTACGCCAAGCTCTAATACGACTCACTATAGGGTATGCTCGGAGCTAGGCATGCTGCA
GACGCGTTACGCATTACGATCCAGAATCCAGAGATTGGAGGTGGCTGGCGTAATATCGGTTTAGTGGGACCTGTG
CCTCCGGGTTCCAGGTGTTGCTAGTGTTTGAACCTCCTGAGCATCATTGGATAACAGTAGCCTCTCACCATGCTCA
TCTTGTGCTTGTATTGGTGGCAGCGGTCCACCATGCCGGTTATGCTGAACTCGGACTCATCACCTTAAATTAACCA
CCTGCCTCAGACTCCGAAACTGCTGGTAGTACAGGCAATCTGCATTCGTCTGCATTCTTCTACAGCCTAGGCTAGC
TATAGACCACACTTGACCACGGCCCGAGCTCCCGGCCGCTTGGATTCTATAGTGTCATATAAAGGCCCGAACAATT
CACTGCACCGTAGTTT

[]> RevE-120m57Ctrl
AAACTACGGTGCAGTGAATTGTTCGGGCCTTTATATGACACTATAGAATCCAAGCGGCCGGGAGCTCGGGCCGTG
GTCAAGTGTGGTCTATAGCTAGCCTAGGCTGTAGAAGAATGCAGACGAATGCAGATTGCCTGTACTACCAGCAGT
TTCGGAGTCTGAGGCAGGTGGTTAATTTAAGGTGATGAGTCCGAGTTCAGCATAACCGGCATGGTGGACCGCTGC
CACCAATACAAGCACAAGATGAGCATGGTGAGAGGCTACTGTTATCCAATGATGCTCAGGAGGTTCAAACACTAG
CAACACCTGGAACCCGGAGGCACAGGTCCCACTAAACCGATATTACGCCAGCCACCTCCAATCTCTGGATTCTGG
ATCGTAATGCGTAACGCGTCTGCAGCATGCCTAGCTCCGAGCATACCCTATAGTGAGTCGTATTAGAGCTTGGCGT
AATCATGGGCATAGCTATT
Sorry. no matches found

5> RevE-119m57Ctrl

CAGCTGACTGCAACCTCTGACTCCTGGGTTCAAGTGACTCTCCTGCCTCAGCCTTCTGAATAGCTGGGATTACGGG
CAAGTACCACCATACCTAGCTAATTTTGTATTTTTAGCAGAGACGGGGTTTCTCCATGTTGGTCAGGCGGTTCTCG
AACTCCCGACCTCAGGTGATCCACCTGCCTCTGCCTCCC

2> RevE-270m50Ctrl
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

5> RevE-169m50Ctrl
CAGCTCTCCACAACCTCCGCCATCGTGGGTTCCAGCAGATTCTCCTGCCTCGGCCTCC CAAGTAGCTGGGAATACA
GGCACGCTCCAATACACCTGGCTAATTATGTATTTTTAGTAGAGACAGGGTTTCTCCATGTTGGTCAACCTGGTCT
GGAACTCCTGACCTCGGGTAATCAACCCACTTCAGCCTCCC

Figure 3 Continued

3> RevE-77m43BD
CAGCTCACTGCAACCACCACCTCCCAGGTTCAAGTGATTATCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGA
TGCCCACCAACACACCAGGCTAATTTTTTGTATTTTTAGTAGAGATGGGGTTTCATCAGGTTGGCCAGGCTGATCT
TAAACTCCTGACCTCAGGTAATCCACCCGCCTCAACCTCCG

2> PK1601mM-13_m37-7+++
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

2> PK1601mM-11_m37-5+++
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTTCC

[]5> PK1601_mM-1_m57-6---------
CAGCTCACTGCAGGCTCCGCCTCCCGGGTTCACGCCATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGACTACAGG
CGCCCACCACCATGCCCAGCTAATTTTTGTATTTTTAGCAGAGACGGGGTTTCACCATGTTGGCCAGGATGGTCTC
CAAACTCCTGACCTCCTGAGACACCTGTGTCGGGGTCCCAAACTGTGGGAGTACAGGCAACTCTGAATTTTTGGAC
AAGACTCTTCGAGCCTATGCTACTATCTACACCACACCGCGTGGGGGCCCCAGCTCGCGGCCGCTGTATTATATAA
TA

3> PK1601mM-60+++
CAGCTCAATGCAACCTACACCTCCTGGGTTCAAGTGATTCTCACGCCTCAGCCTCCTAAGTAACTGGGATTACAGG
GGCGCACCACCACACCTGGCTAATTTTTTGTATTTTTAGCAGAGATGGGCCATGTTGGCCAGGCTGGTCTTGAACT
CCTGACCTCAAGTGATCCACCTGCCTCGGCCTCCC

2> PK1601MM-59+++
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

2> PK1601mM-58+++

Figure 3 Continued

CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

3[]> PK160ImM-57++−
ATCTATGACATGATTGCCCCGATTCTCCAAGCTCTAATTCTACTGAATGTTCGGAACGCTCCATCCACGCATGCCG
TAAACGCTTTACTCCTCGGTTCCAGAATGCGGG
ATTGCCTGTACTTCCATCAGTTAGGGAGGCCAAATCCTACGGATCATATGAGGCTATGAGACCAAGACCCACCTT
ATCAACATGAAGAATCCTGGTCTCTACTAAAAATACAATATTAGCCAGGTTTCATGGTATATGCTTGTAATCCTAG
CTACTCACAAGGCTGAGGCAGAGGAATTACTTGAACCTGTGAGGCGGAGGTTTCGGTGAGCTGAGATTGTCCAAA
CACCCTCCAATCTGAATTCGTTGACAAGCTTTTCGAGCCTAGGCTAGCTCTAGACCACACGTGTGGGGGCCCGAGC
TCGCGGT

> PK160ImM-55+++
ACGTTGCCTGTTCGCAGTTATCGCTACTTGGGAAGTCGTCCCATCTGAGCCGTCGATCGATCCAGAATCGG
ATTGGAGGTGTTGCCAACATTGAGTCACTGCAGCTTTGACCTCCTGAGTGCATGTGGCTTATTCCACCTCAACCTC
CTGAGGAGTTGGGACCACCAGTGTTCAACACCACATCAGGCTAATTTAATATTTTGTAGAAATGAAGACTTACTAT
TATGTCCAGGCTAGTATTAAAATACTGGGGTTAAGCAAGACTCCCCCCTTGTTGTTCCCAAATGCTGGGGCGACAA
CAGGTATTGATTTTTCGACAAGCTTCTTCGAGCCTCCGATGGTTCTATACACCACACGTGGGGCCCGAGCTCTCGC
CGCTG

2> Pk160ImM-54+++
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

2> pk160ImM-53+++
CAGCTCACCGAAACCTGCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

5> pk160ImM-52+++
CAGCTCACTGCAACCTCCGCCTCCTGGATTCAAGCGATTTTCCCGCCTTAGCCTCCTGAGTAACTGGGACTAGAGG
CAGGTACCACCACGCCCAGCTAATTTTTGTATTTTTAGTAGAGACGAGGTTTCACCATGTGGGCCAGGCTGGTCTT
AAACTCCTGACCTCAAGTGATTTGCCCAACTCAGCCTCCC

Figure 3 Continued

5> pk1601mM-51+++
CAGCTCACTGCAACCTCCGCCTCCTGGATTCAAGCGATTTTCCCGCCTTAGCCTCCTGAGTAACTGGGACTACAGG
CAGGTACCACCACGCCCAGCTAATTTTTGTATTTTTAGTAGAGACGAGGTTTCACCATGTGGGCCAGGCTGGTCTT
AAACTCCTGACCTCAAGTGATTTGCCCAACTCAGCCTCCC

2> pk1601mM-50--------
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

8> pk1601mM-49--------
GACTCATTGCAACCTCTGCCTCCTGGGTTTAAGCCGTTCTCATGCCTCAGCCTCCCGACGTAGCTGGGATTATAGG
CATGCGCCACCACCCCCAGCTAATTTTTGTATTATCAGTAGAGATGGGGCTTCGCCATGCTGGCCAGGCTGGTCTT
GAACTCCTGACCTCAAGCAATCCGCCCAACTCGGCCTCCC

pk1601mM-47--------
CAGCTCACCGAAACCTCCGCCTCACGGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

2> pk1601mM-48---------
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

2> pk1601mM-44--------
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGAA
CTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

2> pk1601mM-42---------
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTTGTATGTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

[]10> pk1601mM-37+++

Figure 3 Continued

GACAGGTATGACCATGATTACGCCAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAG
ACGCGTTACGTATGGGATCCAGAATTCGTGATTGGAGGGTGTTTTGCACAATCTCAGCTCACCGCAACCTTTGCCT
CACGGGCTCAAGTGATTCTCATGCTTGATCCTACCAAGTAGCTGGGATTACAGGCACATGCCATCATGCTGAGCTA
ACTTTGGTATTTTTGGTAGAGACGAGGTTTCACCATGTTGGCCAGGCTGTCTCAAACTCCTGACCTCAGATGATCC
GTCCACCTCAGCCTCCC

9> pk1601mM-35+++
CGGCTCACTGCAAGCTCTGCCTCCCGGGTTCATGCCATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGACTGCAGG
TGGCCGTCACCACGCCCGGCTAATTTTTTGTATTTTTAGTAGAGACAGGGTTTCACCATGTTAGCCAGGATGGTCT
CGATCTCCTGACCTCGTGATCTGCCCGCCTCAGCCTCCC

5> pk1601_mM-32+++
CACGTCACTGTAATGTCCATCTCCCGGGTTCAGGTGATTCTCCTGCCCCAGCCTCCTGAGTAGCTGTACAGGCGTG
CACCACCATGCCCGACTAATTTTTGTACTTTTAGTAGAGATTGGGTTTCACCGTGTTGGTCAGGCTGGTCTTGAACT
CCTGACCTCAAGTGATCTGCCTGCCTCAGCCTCCC

4> pk1601_mM-31+++
CAGCTTACTGCAACCTTTGCTTCCCAGTTTCAAGTGATTCTCCTGTCTCATGCTCCAGAGAACCCGGTACTACAGG
CACACGCCACCATGCTCGGCTAATAATTTATGTTCTTAGAATAGAGATTGGTTTTCACCGATTT

6> pk1601_mM-30+++
TGGCTCACTGCAACCTCTGCCACCCGGATTTAAGCAATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGG
CGCCTGCCACTGCTCTGAGCTAATTTTTGTATTTTTGGTAGAGACGGGATTTCACCATCTTGGCCAGGCTGGTTTTA
AACTCCTGACCTCATGATCCACCCGCCTCGGCCTTCC

7> pk1401_mM-24+++
TGGCTTACTGGAACCTTCGCCTTCCGGGTTCAAGAGATTCTTCTGCCTTAACCTTCCGAGAGGCTGGGACTACAGG
CATGCGCCACCATGCCCAGCTAGGTTTTGGATTTTTAAGAGAGATGGGGTTTCCCCATGTTGGCCAGGATGATCTC
GATCTCTTGACCTCGTGATCTGTCCGGCTTAAGACTTCCAAACTGGTGGGAGTACAGGC
AATCTGAATTCGTCGACAAGCTTTTCTAGCCTAGGCTAGCTCTAGACACACGTGTGGGGGCCCGAGCTCGCGGCC
GCTGTA

4> pk1401_mM-23+++
CGGTTCATTGCAACCTCCGCTTCCTAGGGTCCAGTGATCCTCCTGCCTCAGTCCCCCAAGTGGCTGGGACTACAGG
CATGTGCCACCACATCTGGCTAACTTTTGTATATTTAGTAGAAACAGGGTTTCACCATGTTGGCCAGGCTGGTCTC
GAACTCCTGGCCTCAAGTGATCCACCCGCCTTGGCCTCCC

Figure 3 Continued

2> pk1401_mM-22+++
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

7> pk1401_mM-21+++
TGGCTCACTGCAACCTCTGCCTCCTGGGTTCAAGTAATTCTCCTGCCTCAGCCTCCCGAGTACCTGGGACTACAGG
CACCCACCACCACGCTCAGCTAATTTTTGTATTTTTAGTAGAGACGGGGTTTCACCATATTGGCCAGGCTGGTCTC
GAACTCCTGACCTTGTGATCCCCCCGCCTCGGCCGCCC

2> pk1401_mM-20+++

CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

2> pk1401_mM-19+++

CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

2> pk1401_mM-18+++

CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

2> pk1401_mM-17+++

GGGAGGCCAAATCAGATGGATCATCTGAGGTCAGGAGTTCAAGAACCACCTTATCAACATGAAGAATCCTGGTCT
CTACTAAAACTACAAAATTAGCCAGGTATCATGGCAAATGCTTGTCATCCTAGCTACTCAGAAGGCTGAGGCAGA
GGAATCACTTGAACCTGTGAGGCGGAGGTTTCGGTGAGCTG

4> pk1401_mM-16+++

Figure 3 Continued

CAGCTCACTGCAACCTCCCCCTCCTGGGTTCAAGCGATTCTCTTGCCTCAGCCTCCTGAGTAGCTGGGATTACAGG
TGCCCACCACCACGCCCAGTTAATTTTTTGTAGTTTTAGTACAGACGAGGTTCCACTGTGCTGATCAGGCTAGTCT
CGAACTCCTGACCTCAGGTGATCCACCTGCCTTGGCATCTC

2> pk1401_mM-14+++
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

4> pk1401_mM-10--------
CAGCTGACTGCAGTCTTGACCTCGAAGGCTCAAGCGATCCTCCCACCTCTCAGCCTCA CAAGTAGCTGGGACTACT
ACTGACACGCCTCACCACACCCAGCATTTTTTTTTTTTGGTAGAAACAGGGTTTCATTATGTTGCCCAGGGTGGTCT
CAAACTCCTGAGCTCAAGTGATCCTCCCACTCGGCCTCCC

2> pk1401_mM-8--------
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

pk1401_mM-7--------
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGTCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

pk1401_mM-6--------
CAGCTCACCACAACCTCCGCCTCCTGGGTTCCAGCGATTCTCCTGCCTCGGCCTCCCAAGTAGCTGGGATTACAGG
CACGCACCAATACACCTGGCTAATTTTGTATTTTTAGCAGAGACAGGGTTTCTCCATGTTGGTCAACCTGGTCTGT
AACTCCTGACCTCGGGTAATCAACCCACTTCAGCCTCCC

pk1401_mM-5--------
CAGCTCACTGCAACCTCCATTTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCGGAGTAGCTGGGACCACAGA
CGTGTGCCACCATGCCTGGGTAATTTTCATATTTTCAGTAGAGGTGGGGCTTTGCCACATTGTCCAGGCTGGTCTT
GAACTCCTGACCTCAGGTGATCCGCCCGCCTCAGCCTCCC

PK1401_mM-4--------

Figure 3 Continued

TGGCTCACTGCAACCTCCGCCTCCCAGGTTCAAGCAATTCTCCTGCCTCAGTCTCCCGAGTAGCTGGGACTACCGG
CGAGTGCTACCATGCCTGCGTAATTTTTTGTACTTTTAGTAGAGTTGGAGTTTCACTACGTTGGCCAGGCTGGTCTC
AAACTCCTGGCCTCAAGTGATCTGCCGGCCTCAGCCTCCC

pk1401_mM-3-------
CGGCTCACTGCAAGCTCCGCCTCCCGGGTGCACGCCATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGACTACAGG
CGCCCGCCACCACGCCCGGCTAATTTTTTGTATTTTTAGTAGAGGCAGGGTTTCACTGTGTTAGCCAGGATGGTCT
CGATCTCCTGACCTCGTGATCCGCCCGCCTCTGCCTCCC

pk1401_mM-2---------
TGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGCGATTACAGGCATCCGCCACCACACCCAACTAATTTTGTATTTTT
AGTAGAGACAGGTTTTCTCCATGTTGGTCAGGCTAGTCTCGAATTCCTGACCTCAGGTGATCTGCCTGCCTTGGCT
TCCCAAAGTGCTGGGATTACAGGCGTGAGCCACTGTGCCTGGCCAAAGCTATTTC

pk1401_mM-2-------
CAGCTCACTGCAACCTCACCTCCCGGGTTCAAGTGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGCGATTACAGGC
ATCCGCCACCACACCCAACTAATTTTGTATTTTTAGTAGAGACAGGTTTTCTCCATGTTGGTCAGGCTAGTCTCGA
ATTCCTGACCTCAGGTGATCTGCCTGCCTTGGCTTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACTGTGCCTGG
CCAAAGCTATTTCTTTTTTTCTTTTTCCTTTTTTTTTTTTTTTTTTGAGACGGAGTCTCGCTGTGTCCCCCAGGCTGGAG
TACAATGGCATGATCTCGGCTCACTGCAACCTCTGCCTCCCAGGTTTCAAGCGATTTTCCTGCCTCAGCCTCCCGA
GTAGCTGGGATTACAGGCACCCACCACCGTGCCCAGCTAATTTTTGTATCTTTAATAGAGATGGGGTTTCACCATC
TTGGCCAGGCTGGTCTTGAACTCCTGACCTCATGATCCACCCACCTCAGTCTCCCAAACTGCTGGGAGTACAGAAT
CTGAATTC

BDc_m34-4-----BD-----
TGGCTCACTGTAACCTCCACCTCCTGGATTCAAGTGATTCTCCTGCCTCAGCCTCCCACGTAGCTGGGACTACAGG
CACACGACACCGCACCCAGCTCATTTTGTATTTTTAGTAGAGACAGGGTTTCACTATGTTGGCCAGGCTGGTCTCA
AACTTCTGACCTCAGGTGATCCACCCACCTCAGCCTTCC

SZb_m37-10-++
CGGCTCACTGCAGCCTCTACCTCCCATGTTCAAGCCATCCTCCAGTCTCAGCCTCTGGAGTAGTTGGGATTACAGA
TGTGTACCACCTCGCCTGGCTAATTTTTGTATTTTTAGTAGAGATGGGGTTTTGCCATGTTGGCCAGGCTGATCTCA
GATTCCTGATCTCAGGTGATCCACCTGCCTTGGCCTCCC

SZb_m37-9+++

Figure 3 Continued

GGCTCACTGCAGCCTCTACCTCCCATGTTCAAGCCATCCTCCAGTCTCAGCCTCTGGAGTAGTTGGGATTACAGAT
GTGTACCACCTCGCCTGGCTAATTTTTGTATTTTTAGTAGAGATGGGGTTTTGCCATGTTGGCCAGGCTGATCTCAG
ATTCCTGATCTCAGGTGATCCACCTGCCTTGGCCTCCC

SZb_m37-7++-
CGGCTCACTGCAGCCTCTACCTCCCATGTTCAAGCCATCCTCCAGTCTCAGCCTCTGGAGTAGTTGGGATTACAGA
TGTGTACCACCTCGCCTGGCTAATTTTTGTATTTTTAGTAGAGATGGGGTTTTGCCATGTTGGCCAGGCTGATCTCA
GATTCCTGATCTCAGGTGATCCACCTGCCTTGGCCTCCC

SZb_m37-5+++
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

SZb_m37-3+-+
CAGCTATGACCTGATTACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGAC
GCGTTACGTATCGGATCCAGAATTCGTGATT
GCCGGGACTTCGAACCGTCTGGGCTGCCTGAAAGCTTGGACTACCAGGGGTAAGCGGTTCAGGGGCCTCATTATC
AACAGGAACTGTGATGACATGTACTAACAACACTGCCCAGGTCGGGTTTGATGGCAAATGCAGGACATACAAAAT
ACTAATATGGCTGCAGGGCTGGAATCAATCGAACGTGGGAGGGATCCGTCTGCCTGAGCCGACAAAGCTGATGCA
AGTTCCAACATGAATTCGTCGACAAGCTTCTCGAGCCTAGGCTAGCTCTAGACCACACGTGTGGGGGGCC

BDc_m34-10----BD-----
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

SZb_m37-2+++
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

BDc_m34-3----BD-----
TGGCTCACTGTAACCTCCACCTCCTGGATTCAAGTGATTCTCCTGCCTCAGCCTCCCACGTAGCTGGGACTACAGG
CACACGACACCGCACCCAGCTCATTTTTGTATTTTTAGTAGAGACAGGGTTTCACTATGTTGGCCAGGCTGGTCTCA
AACTTCTGACCTCAGGTGATCCACCCACCTCAGCCTTCC

Figure 3 Continued

BDc_m34-1-----BD------
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

pk211201_M39-2----BD-----
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

CtrlC_m57-2----------
TGGCTCACTGCAACCTCCACCTCCCGGGTTCAAGCAATTCTCGTGCCTCAGCCACCTGAGTAGCTGGGATTATAGG
TGTGCGCCACCACACCCGGCTAATTTTTAAATTTTTTGTAGAGACGGGGTTTCACCCTGTTGGCCAGGCTGGCCTC
GAACTCCTAATCTCAGGTGATCTGCCCACCTTGGCCTCCC

BDd_m43-19----BD------
CAGCTGACTGCAACCTCCACTTCCCAGGTTCAAGCGATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGAACTAGAAG
CGTGCACCACCACATCCCGCTAATTGTGTGTGTGTGTGTGTGTTTGTTTAGTAAAGGGGGGGTTTCACCATGTTGG
TCAGGCTGGTCTCGAACTCCTGACAGGTGATCCACCCGCCTTGGCCTCCC

SZc_m37-26+++
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

BDd_m34-19-----BD------
TGGCTCACTGTAACCTCTGCCTCCTGGGTTCAAGTAATTCTCCTGTCTCAGCCTCCTGAGTAGCTAGGATTACTGGT
GCCCGCCACCATGCCCGGCAAATTTTTGTATTTTTAGTAGAGATGGGGTTTCACTATGTTGCCCAGGGTGGTCTCA
AACTCCTGACCTCAAGTGATCCACCTGCTTCAGCTTCCC

BDd_m34-14-----BD------
CAGCCCAGTGCAAGCTCCGCCTCCCAGGTTCACGTCATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGACTACAGG
CGCCCGCCACCACGCCCAGCTAATTTTTTGTATTTTTAGTAGAGACAAGGTTTCACCGTATTAGCCGGGATGGTCG
CTATCTCCTGACCTCGTGATCTGCCCGCCTCGGCCTCTC

Figure 3 Continued

BDd_m43-14-----BD-----
CTCTGCTCACTGCAGCTTCTGCCTCCCGGGTTCAAGTGATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGGACTACA
GGCATGCACCACCACACCCAGCTAATTTTTGTATTTTTAGTAGAGACGGGGTTTCACCATGTTGGCCAGGATGGTC
TCTATCTCTTGACCTCATGATCCGCCCGCCTCAGCCTTCC

SZc_m37-15+++
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

SZc_m37-10+++
CAGCTCACTGCAGGCTCCGCCTCCCGGGTTCACGCCATTCTCCTGCCTCAGCCTCCGCAGTAGCTGGGACTACAGG
CGCCCACCACCATGCCCAGCTAATTTTTGTATTTTTAGCAAAGACAGGGTTTCACCATGTTAGCCAGGATGGTCTC
GATCTCCTGACCTCATGATCCACCTGCCTCGGCCTCCC

SZc_m37-7+++
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

SZc_m37-5+++
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

SZc_m37-3+++
CAGCTCACTGCAGGCTCCGCCTCCCGGGTTCACGCCATTTTCCTGCCTCAGCCTCCCCAGTAGCTGGGACTACAGG
CGCCCATCACCATGCCCAGCTAATTTTTGTATTTTTAGCAAAGACAGGGTTTCACCATGTTAGCCAGGATGGTCTC
GATCTCCTGACCTCCTGATCCACCTGCCTCGGCCTCCC

pk0301_M39-14-----BD-----
AAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

PK0301_M37-14+++

Figure 3 Continued

CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

PK0301_M37-11+++
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTATGCCTTAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGCGGTTCTTGA
ACTCCTGACCTCACATGATCCATTTGATTTGGCCTCC

RevCompSZB_M37-6+++
CAGCTCACTGGCAGTCTCAATCTTCCAAGTTCAAGGTGATTATCCCATCTCAGCCTCCCGAGTAGCTGAAACTACA
GGTGCATACTACCACGCCTAGCTAATTTTTTTTTGTAGAGATGGGGTTTTGGCCATGTTGCCCAGGCTGCTCTCGA
ACTTCTGGGCACAAGTGGTCCACCCACCTTGGCCTCCC

RevCompPK1401_mM-17+++
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTAGTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

RevCompPK1601mM-33+++

CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

RevCompPK1601mM-39+++
CAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTCTGAGTAGCTAGGATGACAAG
CATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCATGTTGATAAGGTGGTTCTTGA
ACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCC

CUTPK1601_mM-1_m57-6---------

GAACCACCATTACGCCAACTCTAATACGACTCACTATAGGGAAAGCTCGGTACCACGCATGCTGCAGACGCGTTA
CGTATCGGATCCAGAATTCGGGATTGGAGGGTGTTTGCACAATCTCAGCTCACTGCAGGCTCCGCCTCCCGGGTTC
ACGCCATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGACTACAGGCGCCCACCACCATGCCCAGCTAATTTTTGTA

Figure 3 Continued

603 6818 .060602

TTTTTAGCAGAGACGGGGTTTCACCATGTTGGCCAGGATGGTCTCCAAACTCCTGACCTCCTGAGACACCTGTGTC
GGGGTCCCAAACTGTGGGAGTACAGGCAACTCTGAATTTTTGGACAAGACTCTTCGAGCCTATGCTACTATCTACA
CCACACCGCGTGGGGGCCCCAGCTCGCGGCCGCTGTATTATATAATA

CUTPK1601mM-57+++
CATCTATGACATGATTGCCCCGATTCTCCAAGCTCTAATTCTACTGAATGTTCGGAACGCTCCATCCACGCATGCC
GTAAACGCTTTACTCCTCGGTTCCAGAATGCCGGGATTGCCTGTACTTCCATCAGTTAGGGAGGCCAAATCCTACGG
ATCATATGAGGCTATGAGACCAAGACCCACCTTATCAACATGAAGAATCCTGGTCTCTACTAAAAATACAATATT
AGCCAGGTTTCATGGTATATGCTTGTAATCCTAGCTACTCACAAGGCTGAGGCAGAGGAATTACTTGAACCTGTGA
GGCGGAGGTTTCGGTGAGCTGAGATTGTCCAAACACCCTCCAATCTGAATTCGTTGACAAGCTTTTCGAGCCTAGG
CTAGCTCTAGACCACACGTGTGGGGGCCCGAGCTCGCGGT

CUTPK1601mM-55+++
ACGTTGCCTGTTCGCAGTTATCGCTACTTGGGAAGTCGTCCCATCTGAGCCGTCGATCGATCCAGAATCGGATTGG
AGGTGTTGCCAACATTGAGTCACTGCAGCTTTGACCTCCTGAGTGCATGTGGCTTATTCCACCTCAACCTCCTGAG
GAGTTGGGACCACCAGTGTTCAACACCACATCAGGCTAATTTAATATTTTGTAGAAATGAAGACTTACTATTATGT
CCAGGCTAGTATTAAAATACTGGGGTTAAGCAAGACTCCCCCCTTGTTGTTCCCAAATGCTGGGGGGACAACAGG
TATTGATTTTTCGACAAGCTTCTTCGAGCCTCCGATGGTTCTATACACCACACGTGGGGCCCGAGCTCTCGCCGCT
G

utPK1601mM-39+++
GGGAGGCCAAATCAGATGGATCATCTGAGGTCAGGAGTTCAAGAACCACCTTATCAACATGAAGAATCCTGGTCT
CTACTAAAAATACAAAATTAGCCAGGTATCATGGCAAATGCTTGTCATCCTAGCTACTCAGAAGGCTGAGGCAGA
GGAATCACTTGAACCTGTGAGGCGGAGGTTTCGGTGAGCTGA

CutPK1601mM-37+++
GGGAGGGTGTTTTGCACAATCTCAGCTCACCGCAACCTTTGCCTCACGGGCTCAAGTGATTCTCATGCTTGATCCT
ACCAAGTAGCTGGGATTACAGGCACATGCCATCATGCTGAGCTAACTTTGGTATTTTTGGTAGAGACGAGGTTTCA
CCATGTTGGCCAGGCTGTCTCAAACTCCTGACCTCAGATGATCCGTCCACCTCAGCCTC

CutPK1601mM-33+++
TGGGAGGCCAAATCAGATGGATCATCTGAGGTCAGGAGTTCAAGAACCACCTTATCAACATGAAGAATCCTGGTC
TCTACTAAAAATACAAAATTAGCCAGGTATCATGGCAAATGCTTGTCATCCTAGCTACTCAGAAGGCTGAGGCAG
AGGAATCACTTGAACCTGTGAGGCGGAGGTTTCGGTGAGCTGA

CutPK1601_mM-31+++

Figure 3 Continued

TCAGCTTACTGCAACCTTTGCTTCCCAGTTTCAAGTGATTCTCCTGTCTCATGCTCCAGAGAACCCGGTACTACAGG
CACACGCCACCATGCTCGGCTAATAATTTATGTTCTTAGAATAGAGATTGGTTTTCACCGATTT

CutPK1401_mM-17+++
TGGGAGGCCAAATCAGATGGATCATCTGAGGTCAGGAGTTCAAGAACCACCTTATCAACATGAAGAATCCTGGTC
TCTACTAAAACTACAAAATTAGCCAGGTATCATGGCAAATGCTTGTCATCCTAGCTACTCAGAAGGCTGAGGCAG
AGGAATCACTTGAACCTGTGAGGCGGAGGTTTCGGTGAGCTGA

CutPK1401_mM-2_1++
TCAGCTCACTGCAACCTCACCTCCCGGGTTCAAGTGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGCGATTACAGG
CATCCGCCACCACACCCAACTAATTTTGTATTTTTAGTAGAGACAGGTTTTCTCCATGTTGGTCAGGCTAGTCTCGA
ATTCCTGACCTCAGGTGATCTGCCTGCCTTGGCTTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACT

CutPK1401_mM-2_2+++
GAGACGGAGTCTCGCTGTGTCCCCCAGGCTGGAGTACAATGGCATGATCTCGGCTCACTGCAACCTCTGCCTCCCA
GGTTTCAAGCGATTTTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCACCCACCACCGTGCCCAGCTAATT
TTTGTATCTTTAATAGAGATGGGGTTTCACCATCTTGGCCAGGCTGGTCTTGAACTCCTGACCTCATGATCCACCCA
CCTCAGTCTC

CutSZB_M37-6+++
TGGGAGGCCAAGGTGGGTGGACCACTTGTGCCCAGAAGTTCGAGAGCAGCCTGGGCAACATGGCCAAAACCCCA
TCTCTACAAAAAAAAATTAGCTAGGCGTGGTAGTATGCACCTGTAGTTTCAGCTACTCGGGAGGCTGAGATGGGA
TAATCACCTTGAACTTGGAAGATTGAGACTGCCAGTGAGCTGA

CutSZB_M37-3++
TGCCGGGACTTCGAACCGTCTGGGCTGCCTGAAAGCTTGGACTACCAGGGGTAAGCGGTTCAGGGGCCTCATTAT
CAACAGGAACTGTGATGACATGTACTAACAACACTGCCCAGGTCGGGTTTGATGGCAAATGCAGGACATACAAAA
TACTAATATGGCTGCAGGGCTGGAATCAATCGAACGTGG

PK37-9RfWithM13R
GCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCG
TAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCATTCGCCATTCAGGCTGCGCAACTGTTGG
GGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAG
TTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTGTAATACGACTCACTATA
GGGCGAATTGGGCCCTCTAGATGCATGCTCGAGCGGCCGCCAGTGTGATGGATATCTGCAGAATTCGGCTTGCCT
GTACTCCCAGCAGTTT

Figure 3 Continued

>PK39-4RfWithM13R
CCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGG
GCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGT
AACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTGTAATACGACTCACTATAGGGCGA
ATTGGGCCCTCTAGATGCATGCTCGAGCGGCCGCCAGTGTGATGGATATCTGCAGAATTCGGCTTGCCTGTACTCC
CAGCAGTTT

>PK37-9RrWithM13R
GCCTGTACTCCCAGCAGTTTGAGAGGCCAAGATGGGTGGATCACTTGAGGTCTAGAGCTCAAGACCAGCCTGGCG
ACATGGTGAAACCCCATCTCTACTAAAAATATAAAAATCAGCCAGGTGTGGTGGTGGGCACCTGTAACCCCAGCT
ACTCAGGAGGCTGAGGAAGCCGAATTCCAGCACACTGGCGGCCGTTACTAGTGGATCCGAGCTCGGTACCAAGCT
TGGCGTAATCATGGTCATAGCTGTT

2>PK39-4RrWithM13R
GCCTGTACTCCCAGCAGTTTGAGAGGCCAAATCAGATGGATCATCTGAGGTCAGGAGTTCAAGAACCACCTTATC
AACATGAAGAATCCTGGTCTCTACTAAAAATACAAAATTAGCCAGGTATCATGGCAAATGCTTGTCATCCTAGCT
ACTCAGAAGGCTGAGGCAGAGGAATCACTTGAACCTGTGAGGCGGAGGTTTCGGTGAGCTGAGATTGTGCAAAC
ACCAAGCCGAATTCCAGCACACTGGCGGCCGTTACTAGTGGATCCGAGCTCGGTACCAAGCTTGGCGTAATCAGG
TCATAGCTGTT

>PK34-6rwithM13R
GCCTGTACTCCCAGCAGTTTTGAGAGGTCAAGGAAGGAGGATCAGTTGAGTCCGGGAGTTTGAGATGAGCCTGGG
CAACATGGCAAAACCTCGTCTCTACAAAAAATACAAAAAAAGTAAGCCGGGCATGGTGGAGAGGCTATTCGGCT
ATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTC
TTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCA
CGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAG
TGCCGGGGCAGGATCTCCTGTCATCCCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCG
GCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCG
GATGGAAGCCGGTCTTGTCGATCA

>PK37-1withM13R
AACAGCTATGACCTGATTACGCCAAGCTTGGTACCGAGCTCGGATCCACTAGTAACGGCCGCCAGTGTGCTGGAA
TTCGGCTTGCCTGTACTCCCAGCAGTTTGGGAGGCCAAATCAGATGGATCATCTGAGGTCAGGAGTTCAAGAACC
ACCTTATCAACATGAAGAATCCTGGTCTCTACTAAAAATACAAAATTAGCCAGGTATCATGGCAAATGCTTGTCAT
CCTAGCTACTCAGAAGGCTGAGGCAGAGGAATCACTTGAACCTGTGAGGCGGAGGTTTCGGTGAGCTGAGATTGT

Figure 3 Continued

GCAAACACCCTCCAAGCCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCC
AATTCGCCCTATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCG
TTCCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCC
CTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGT
G

>PK37-1rwithM13R
GCCTGTACTCCCAGCAGTTTGGGAGGCCAAATCAGATGGATCATCTGAGGTCAGGAGTTCAAGAACCACCTTATC
AACATGAAGAATCCTGGTCTCTACTAAAAATACAAAATTAGCCAGGTATCATGGCAAATGCTTGTCATCCTAGCT
ACTCAGAAGGCTGAGGCAGAGGAATCACTTGAACCTGTGAGGCGGAGGTTTCGGTGAGCTGAGATTGTGCAAAC
ACCCTCCAAGCCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTCG
CCCTATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTCCCA
ACTTAATCGCCTTGCAGCACATCCCCCTTTCGCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCC
AACAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTG

>PK34-2withM13R
GCCTGTACTCCCAGCAGTTTGGGAGGCCGAGGCGGGCAGATTGCCTGAGCTCAGGAGTTCGAAACCAGCCTGGAC
AACACGGTGAAACCCTGTCTCTACTAAAAATACAAAAAATTAGCCAGACGTGGTGGTGCATGCCTGTAGTCCCAG
CTAGTCAGGAGGCTGAGGCAGGAGAATCACTTGAACCCAGCAGGAAAAGGTTGTGGTGAGCTGAGATTGTGCAA
ACACCCTCCAAGCCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATT
CGCCCTATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTAC
CCAACTTAATCGCCTTGCAGCACATTCCCCTTTCGCCAGCTGGCGTAATAGCTAAGAGGCCCGCACCGATCGTCCC
TTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTG
GTTAC

>PK34-7withM13R
GGAGGGTGTTTGCACAATCTCGGCTTACTGCAACCTCCACTCCTGGGCTTAAACGGTCCTCCCACCTCATCTTCCC
GAGTAGCAGGGTCCACAGGTGCACACCACCATGCCTGGCTATATTTTTTTTTTTTTTTGGATTTTTGATAAAGACAG
GATGTCAACATGTTGCCCACGCTGGTCTTCAACCCCTTGAACTCAAATTCATCTGCTTCTGCCTCCCAAACTGGTG
GGAGTCTTGAGGTGGGCGAACCACCTGATGTTACGAATATGAGACTTTTCGGCCTGATTCCGGCCAAACTCTCGTC
TTATTTTTTATAATCTAATAAATCCCATCTAGGGGCTAGGGT

>PK34-8withM13R
GGAGGGTGTTTGCACAATCTCAGCTCACCGAAACCTCCGCCTCACAGGNTCAAGTGATTCCTCTGCCTCAGCCTTC
TGAGTAGCTAGGATGACAAGCATTTGCCATGATACCTGGCTAATTTTGTACTTTTAGTAGAGACCAGGATTCTTCA
TGTTGATAAGGTGGTTCTTGAACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCCAAACTGCTGGGAGTACA

Figure 3 Continued

GGCAAGCCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTCGCCCT
ATAGTGAGTCGTATTACAATTCACTGGCCGGCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACT
TAATCGCCTTGCAGCACATC

>PK34-9withM13R
GCCTGTACTCCCAGCAGTTTGGGAGGTCAAGGTGGAGAGATCACTTGAGGTCAGGAGTTCGAGACCAGCCTAACC
AATATGATGAAACCCCATCTCTACTAAAAATACAAAAATTAGCCGGGCGTGGTGGTGCGCACCTGTAATCCCAGC
TACTCAGGAGGCTGAGGCAGGAGAATTGCTTGAACCAGGGAGTCGGAGGTTGCAGTAAGCCAAGATTGTGCAAA
CACCCTCCAAGCCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTC
GCCCTATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACC
CAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAA

>PK37-3 1withM13R
CCTGTACTCCCAGCAGTTTGGAAGTGGATCACTTGAGGCCAGGGACTCAAGACCAACCTGGCCAATATGGCAAAA
CCCGGCTAAAAATACAAAAATTAGCTGGACATGGTTGCAGGTGCCTGTAATCCCAGCTACTCGGGAGGTTGTGGC
ATGAGAATCACTTGAACCTGGGAGGCAGAGGCTGCAGCGAGCAGAGATTGTGCAAACACCCTAAGCCGAATTCT
GCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTCGCCCCTATAGTGAGTCGCAT
TACAATTTACTGGCCCGTCGTTTTACAACCGTCCCGACTGGGAAAACCCTGGCGTTAC

>PK37-7withM13R
GCCTGTACTCCCAGCAGTTTGGGAGGCCAAATCAGATGGATCATCTGAGGTCAGGAGTTCAAGAACCACCTTATC
AACATGAAGAATCCTGGTCTCTACTAAAAATACAAAATTAGCCAGGTATCATGGCAAATGCTTGTCATCCTAGCT
ACTCAGAAGGCTGAGGCAGAGGAATCACTTGAACCTGTGAGGCGGAGGTTTCGGTGAGCTGAGATTGTGCAAAC
ACCCTCCAAGCCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTCG
CCCTATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCC
AACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTC
CCAACAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGT

>PK39-2withM13R
GCCTGTACTCCCAGCAGTTTGGGAGGCTGAGGCGGTTGGATCACAAGGTTAGGAGTTTGAGGCCAGCCTGGCCAA
TAAGATGAAACCCCATCTGTACTAAAAATACAAAAATTAGCCAAACGTGGTGGTGGGCACCTGTAGTCCCAGCTA
CTTGGGAGGCTGAGGCAAAAAAATTGCTTGAACCTGGGAGGCGGAGGTTGCAGCGAGCTGAGATTGTGCAAACA
CCCTCCAAGCCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTCGC
CCTATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCA
ACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCC

Figure 3 Continued

AACAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAGCCCCGGCGGGTGTGGTGGTTA
CGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCC

5>BD43-13
GCCTGTACTCCCAGCAGTTTGGGAGGCCGAGGTGGGCGGATGGCCTGAAGCCAGGAGTTTGAGACTAGCCTGGCC
TACATGGTGAAAACCTGTCTCTACTAAAAATACAATAATTAGCCGGACATGGTGACACCTATAATACCAGCTACT
CGGGAAGCTGAGCCATGAGAATTGCTTGAACCCGGAAGGTGGAGGTTGCAGTGAGCTGAGATTGTGCAAACACC
CTCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGC
GAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCT
TGACGAGTTCTTCTGAATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTGC
GGCATTTTGCCTTCCTGTTTTGCTCACCCACAAACCCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGG

BD43-18withM13R
GCCTGTACTCCCAGCAGTTTGGGAGGCCAAAGCGGACGGATCATATGAGGTCGAGAGTTCAAGAACCATGTTATC
AATGTGAAAAATCTGGGTCTATACTAAAAACACAAATTTACCCAGGGTTGATGGAAGATGCTGGTCATCCTAATT
CCTCAGAAGGCTGAGGCAGAGGAATCATTTGAACCTGGGAGGCGGACGTTCAGGGGACCTGAAATGGGGCAACC
ACCTTCAAAGCCGAATTTTGCAAATTTCCATAACATGGGGGGCGCGTTCAACCTTGCTTTTAAAGGGCCCATTTCC
CTTATATGGAGTCGATTTACAATTAACGGGCGGTCGTTTTACACCTTTGGATGGGAAAAACCCTGCGTACCCCA

>Ctrlm57-7withM13R
ACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACC
TGTCCGGTGCCCTGAATGAACTGCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCG
CAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCC
TGTCATCCCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCC
GGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGT
CGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGCAT
GCCCGACGGCAGGATCTCGTCGTGACCATGGCGATGCCTGCTTGCCA

>pk50-26withM13R(-46)
TTAAAACCGAAATGCCATGATACGCCAAGCTTGGTACCGAGCTACGGACCCACTAGCTAACGGCCGCCAGTGTGC
CTGACCTCTTATCCCTGCACGATATCCACTCACACTGCTGGCTGTCCGTGCATGCATCTACCGGGCTCAATTCGCCC
TATAGTGAGTCGGATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGTGTTACCCAAC
TTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTTGGCGCAATAGCGAAGAGGCATCGCTCGATCGCCCTTTC
CAACAGCTTGCGCAGCCAGAATGGCTAATGGACGCGCCCTGTCTCCGGCCGCATTAATCCGCGGCGGGTGTGGCG
GTTACCCCGCAGCAGTG

Figure 3 Continued

>PK34-1withM13R
GGAGGGTGTTTGCACAATCTGGAGGGTGTTTGCACAATCTCGGCTCACCACAACCTCTACCTCCCAGGTTCAAGCA
ATTCTGCCTCAGCCTCCCAAGTAGCTGGGACTACAGGCGTGCACCACCACACCTGGCTAATTTCTGTATTTTTAGT
AGAAACAGGGTTTCACCATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTTGTGATCCGCCTACCTTGGCTTTCCA
AACTGCTGGGAGTACAGGCAAGCCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGA
GGGCCCAATCCGCCCTATAGTGAGTCGTATTACAATCCACTGGCCGAAGTTTACAACGGCGTGACTGGGAAAACC
CTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGAAATAGCGAAGAGGCCCGCA
CCGATCGCCCTTCCC

>PK34-3withM13R
GGAGGGTGTTTGCACAATCTCTGCTCACTACAACTTCTACCTCCCAGGCTCAAGCAATCCTCCCATGTAGCTGGGA
CCACAGGTGTGCACCACCATGCCAAGCTAATTTTTGTATTTTTTTGTAGAGTGAGGTTTCACCATATTGCCCAGGTT
GGTCTTGAACTCCTAAGCTCAAGCAATCCACCTGCCTCAGCTTCTCAAACTGCTGGGAGTACAGGCAAGCCGAATT
CTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTCGCCCTATAGTGAGTCGTAT
TACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAG
CACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAAAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCT
GAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACG

>PK34-4withM13R
GGAGGGTGTTTGCACAATCTCGGCTCATGGCACCCCTCGCCTCCCAGATTCAAATGATACTCCTGCCTCAGCCTCC
TGAGTAGCTGGGATTACATGCATGCGCCACCATGCCCAGCTAATTTTTTGTATTTTTAGTAGAGACGGGGTTTCAC
CATGTTGGCCAGACTAGACTTGAACTCCTGACCTCGTGATCCACCCACCTCAACCTCCCAAACTGCTGGGAGTACA
GGCAAGCCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTCGCCCT
ATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACT
TAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAAAGGCCCGCACCGATCGCCCTTCCCAA
CAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACG

>PK34-5withM13R
GGAGGGTGTTTGCACAATCTCAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTC
TGAGTAGCTAGGATGACAAGCATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCA
TGTTGATAAGGTGGGTCTTGAACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCCAAACTGCTGGGAGTACA
GGCAAGCCGAATTCTGCAAATATCCATCACACTGGCGGCCGTTCGAGCATGCATCTAAAGGGCCCAATTCGCCCT
ATAGGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAAC
TTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCC

>PK37-1withM13R

Figure 3 Continued

GCCTGTACTCCCAGCAGTTTGGGAGGCCAAATCAGATGGATCATCTGAGGTCAGGAGTTCAAGAACCACCTTATC
AACATGAAGAATCCTGGTCTCTACTAAAAATACAAAATTAGCCAGGTATCATGGCAAATGCTTGTCATCCTAGCT
ACTCAGAAGGCTGAGGCAGAGGAATCACTTGAACCTGTGAGGCGGAGGTTTCGGTGAGCTGAGATTGTGCAAAC
ACCCTCCAAGCCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTCG
CCCTATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTCCCA
ACTTAATCGCCTTGCAGCACATCCCCCTTTCGCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCC
AACAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTG

>PK37-2withM13R
GGAGGGTGTTTGCACAATCTCAGCTCATTGCAACTTCCAGCTCCCAGGTTCAAGCGATTCTCCTTCCTCAGCCTCC
CAAGTAGTTGGGATTACAGGCATGCACCATCATGCCCGGCTAATTTTTGTATTTTTAGTAGAGACAGGGTTTCACC
ATATTGGCCAGGCTGGTCTTGAACTCCTGACCTCGTGTTCCACCCACCTCAGCCTCCCAAACTGCTGGGAGTACAG
GCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTCGCCCTATAGTG
AGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCG
CCTTGCAGCACATTCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGC
GCAGCCTGAATGGCGAATGGACGCGCC

>PK37-4withM13R
GGAGGGTGTTTGCACAATCTCAGCTCATTGCAACCTCCCAGGTTCAAGCGATTCTCCTGCCTCAGCCTCCTGAGTA
GCTGGGATCACAGGTGTGTGCCACCATTCCTGGCTAATTTTTGTATTTCTAGTAGAGATGGGGTTTTACCATGTTG
GTCAGGCTGGTCTCAAACTCCTGACCTCATGATCTGCCCACCTTGGCCTCCCAAACTGCTGGGAGTACAGGCAAGC
CGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTCGCCCTATAGTGA
GTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGC
CTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTTCCCAACAGTTG
CGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGTCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGC

>PK37-5withM13R
GGAGGGTGTTTGCACAATCTCAGCTCACTACAACCTCTGCCTCCCAGGTTCAAGCGATTCTCATGCCTCGGCTTCT
CAAGTTGCTGGGACTACGGGCACACGCCAGCACGGCTGGCTAATTTTTGTATTTTTAGTAGAGACAGGGTTTCACC
GTCTTGGCCATGCTGGTCTCAAACTCCTGACCTCATGATCCACCCGCCTTGGCCTCCCAAACTGCTGGGAGTACAG
GCAAGCCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTCGCCCTA
TAGTGAGTCGTATTACAATTTACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCCTGGCGTTACCCAACT
TAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAA
CAGTTGCGCC

>PK37-6withM13R

Figure 3 Continued

GGAGGGTGTTTGCACAATCTCAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTC
TGAGTAGCTAGGATGACAAGCATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCA
TGTTGATAAGGTGGTTCTTGAACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCCAAACTGCTGGGAGTACA
GGCAAGCCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTCGCCCT
ATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACT
TAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAA
CAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGC
G

>PK37-8withM13R
GGAGGGTGTTTGCACAATCTTTGCTCACTGCAATCTCCACCTCCCGGGTTCAAGTGATTCTCCTGCCTCAGACTGCT
GAATACTTGGGATTACAGGCACCCGCCACCACACCTTGCTAATTTTTTGGATTTTTAATAGAGATGGGGGTTCACC
ATGTCAACCAGGCTGGTCTTGAACTCCTGACCTTAGGTGATCCACCCACCTCAGCCTCCCAAACTGCTGGGAGTAC
AGGCAAGCCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTCGCCC
TATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAAC
TTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCA
ACAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGG

>PK37-9withM13R
AACAGCTATGACCATGATTACGCCAAGCTTGGTACCGAGCTCGGATCCACTAGTAACGGCCGCCAGTGTGCTGGA
ATTCGGCTTCCTCAGCCTCCTGAGTAGCTGGGGTTACAGGTGCCCACCACCACACCTGGCTGATTTTTATATTTTTA
GTAGAGATGGGGTTTCACCATGTCGCCAGGCTGGTCTTGAGCTCTAGACCTCAAGTGATCCACCCATCTTGGCCTC
TCAAACTGCTGGGAGTACAGGCAAGCCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCT
AGAGGGCCCAATTCGCCCTATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAA
ACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCG
CACCGATCGCCCTTCCCCAACAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAGCGC
GGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTC
CCTTCCTTTCTCGC

>PK37-26withM13R
GGAGGGTGTTTGCACAATCTCGGCTCACAGTAGCCTCTGCCTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCC
CGAGTAGCTGGGATTACAGGCATGCGCCACCATGTCCATCTAATTTTGTATTTTTAGTAGAGATGGGGTTTCTCCA
TGTTGGTCAGGCTGGTCTCGAACTCCCAACCTCAGGTGATCCACCCGCCTTGGCCTCCCAAACTGCTGGGAGTACA
GGCAAGCCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTCGCCCT
ATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGCGACTGGGAAAACCCTGTCGTTACCCAACT
CAATCGCCTTGCAGC

Figure 3 Continued

>PK39-3withM13R
GGAGGGTGTTTGCACAATCTTGGCTCACTGCAACCTCTGCCTCCTGGGCCCAAGCCATCTTCCTACCTCAGCTTCC
CGAGTAGCTGGACTACAGGTGTGAGCCATCACGCCCAGCCAATTTTTGTATTTTTAGTAGAGACGAGGTTTCACCA
TGTTGGCCTGGCTGGCCTTGATCTCCTGACCTAGTGATCTCCCCGCCTCAGCCTCTCAAACTGCTGGGAGTACAGG
CAAGCCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTCGCCCTAT
AGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTA
ATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCAACA
GTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAACGCGGCGGGTGTGGTGGTTACGCGC
AGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCG

>PK39-4withM13R
AACAGCTATGACCTGATTACGCCAAGCTTGGTACCGAGCTCGGATCCACTAGTAACGGCCGCCAGTGTGCTGGAA
TTCGGCTTGGTGTTTGCACAATCTCAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGC
CTTCTGAGTAGCTAGGATGACAAGCATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTC
TTCATGTTGATAAGGTGGTTCTTGAACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCTCAAACTGCTGGGAG
TACAGGCAAGCCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTCG
CCCTATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCC
AACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTC
CCAACAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGG

>PK39-6withM13R
GGAGGGTGTTTGCACAATCTCAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTC
TGAGTAGCTAGGATGACAAGCATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGATGGGGTTTTGCCA
TGTTGGCCAGGCTGGTCTCAAACTCCTGACCTCAAGTGATCCCCCACCTCGGCCTCCCAAACTGCTGGGAGTACAG
GCAAGCCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTCGCCCTA
TAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTT
AATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAAC
AGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTAC

>PK39-7withM13R
GGAGGGTGTTTGCACAATCTGGAGGGTGTTTGCACAATCTCGGCTCACCACAATCTTTGCCTTTCGGGTTCAAGGG
ATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCATGTGCCACCACACCCGGCTAATGTTGTAGTTTTA
GTAGAGACGGGGTTTCTCTATGTTGGTTAGGCTGGTCTCAAACTCCTGACCTCAGGTGATCTACCCGCCTCGGCCT
CTCAAACTGCTGGGAGTACAGGCAAGCCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATC
TAGAGGGCCCAATTCGCCCTATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAA

Figure 3 Continued

AACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCC
GCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTA

>PK39-8withM13R
GGAGGGTGTTTGCACAATCTCAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTC
TGAGTAGCTAGGATGACAAGCATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGATGGGGTTTTGCCA
TGTTGGCCAGGCTGGTCTCAAACTCCTGACCTCAAGTGATCCCCCACCTCGGCCTCCCAAACTGCTGGGAGTACAG
GCAAGCCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTCGCCCTA
TAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTT
AATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAAC
AGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTG

>PK39-9withM13R
GGAGGGTGTTTGCACAATCTCAGCTCATTGCAACCTCCACCTCCCGGGTTCAAGCAATTCCCCTGCCTCAGCCTCC
TGAGTAGCTGGAACTACAGGCACGCGCCACCACGTCTGGTTAATTTTTTTGTATTTTTATAGAGATGGGGTTTTAC
CATGTTGCCCAGGCTGGTCTTAAACTCCTGGGCTCAAGCTATCCACTCGCCTTGGCCTCCCAAACTGCTGGGAGTA
CAGGCAAGCCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTCGCC
CTATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAA
CTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCA
ACAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACG
CGCAGCGTG

>PK39-10withM13R
GGAGGGTGTTTGCACAATCTCAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTC
TGAGTAGCTAGGATGACAAGCATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCA
TGTTGATAAGGTGGTTCTTGAACTCCTGACCTCAGATGATCCATTTGATTTGGCCTCCCAAACTGCTGGGAGTACA
GGCAAGCCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTCGCCCT
ATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACT
TAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAA
CAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGC
GCAGCGTGACC

>PK39-12withM13R
GGAGGGTGTTTGCACAATCTTGGCTCACTGCAACTTTTGCCTCCTGGGTTCAAGCAATTCTCCTGCCTCAGCCTCCC
GAGTAGCTGGGACTATAGGCACGCGCCATCACGCCGGGTTATTTTGTATTTTTAGTACAGACGGGGTGTTTACATG
GTGGTCAAGCTGGGTTTGAACTTCTGACCTCAAGTGATCCTGCCCGCCTCGGCTTTCCAAACTGCTGGGAGTACAT

Figure 3 Continued

GGCAAGCCCGAATTCTGCAGATATCCATCACACCTGGCGGCCGCTCGAGCTTGCATCTAGAGGGCCCAATTCCGC
CCTATTCTGAGTCGTATCTACAA

>BD43-1withM13R
GGAGGGTGTTTGCACAATCTCAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTC
TGAGTAGCTAGGATGACAAGCATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCA
TGTTGATAAGGTGGTTCTTGAACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCCAAACTGCTGGGAGTACA
GGCAAGCCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTCGCCCT
ATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAAC
TTAATCGCCTTGCAGCACATCCCCCCTTTCGCACCTG

>BD43-2withM13R
GGAGGGTGTTTGCACAATCTCAGCTCACCGAAACCTCCGCCTCACAGGTTCAAGTGATTCCTCTGCCTCAGCCTTC
TGAGTAGCTAGGATGACAAGCATTTGCCATGATACCTGGCTAATTTTGTATTTTTAGTAGAGACCAGGATTCTTCA
TGTTGATAAGGTGGTTCTTGAACTCCTGACCTCAGATGATCCATCTGATTTGGCCTCCCAAACTGCTGGGAGTACA
GGCAAGCCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTCGCCCT
ATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACT
TAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAA
CAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGC
GCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCC

>BD43-6withM13R
GGAGGGTGTTTGCACAATCTGGGGTTCAAGGGAAGAGTCCAGGCTGCAGATAAAGATTTGGGAGTTGTCAGTATA
GCAATTTCATTGTTGTTATTACTGTTGTTGTTTTGTAGAGATAGGGTCTCACTATGTTGCCCACGCTGGTCTTGAAC
TCCTGAGCTCAAGCGATCCTCCTGCTTCAGCCTCCCAAACTGCTGGGAGTACAGGCAAGCCGAATTCTGCAGATAT
CCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTCGCCCTATAGTGAGTCGTATTACAATTCAC
TGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCC
TTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCAACAGTTGCGCAGCCTGAATGGCGAA
TGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCA
GCGCCCTAGCGCCCG

4>BD43-8
GGAGGGTGTTTGCACAATCTTGGCTCACTGCAACCTCCACCTCGCAGTTCAAGCAATTCTTGTGCCTTAGCCTCCT
GAATAGTAGCTGGGATTACGGGCGTGTGCCATCACACCCAGCTAATTTTTGTATTTTTAGTAGAGACAGTTGTCCA
GGCTGGTCTTGAATTCCTGGCCTCAAGAGATCCGCTGGCTTTGGCCTCTCAAACTGCTGGGAGTACAGGCAAGCCG
AATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTCGCCCTATAGTGAGTC

Figure 3 Continued

GTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTT
GCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGC
AGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGAC
C

4>BD43-8(2)withM13R BD43-8 (178, 100, 11q22.3)
GGAGGGTGTTTGCACAATCTTGGCTCACTGCAACCTCCACCTCGCAGTTCAAGCAATTCTTGTGCCTTAGCCTCCT
GAATAGTAGCTGGGATTACGGGCGTGTGCCATCACACCCAGCTAATTTTTGTATTTTTAGTAGAGACAGTTGTCCA
GGCTGGTCTTGAATTCCTGGCCTCAAGAGATCCGCTGGCTTTGGCCTCTCAAACTGCTGGGAGTACAGGCAAGCCG
AATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTCGCCCTATAGTGAGTC
GTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTT
GCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGC
AGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGAC
C

>BD43-9withM13R
GGAGGGTGTTTGCACAATCTCAGCTCACTGCAACCTTCGCCTCCCGGGTTCAAGTGATTCTCCTGCCTCAGCCTCC
TGAGTAGCTAGGACTATAGATGCCCCCACCACGCCTGGCTAATATTTGTATTTTTTTAGTACAGTCGGGGTTTTGC
CATGTTGGCCAGGCTGATCTCGAACCCCTGACCTCAACTGATCCACCCACCTCGGCCTTCCAAACTGCTGGGAGTA
CAGGCAAGCCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTCGCC
CTATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAA
CTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCA
ACAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAGCCCGGCGGGTGTGGTGGTTAC

>BD43-10withM13R
GGAGGGTGTTTGCACAATCTCAGCTCACTGCAACCTCCCTCTTCTGCATTCAAATGATTCTCATGCCTCAGCCTTCC
GAGTAGCTGGAATTACAGACATGTACTACCACACCAGGCTAAGTTTTGTATTTTTAGTAGAGACGAGGTTTCACCA
TGTTGGCCAGGCTGGTCTTGAACTCCTGGCCTCAAGTGATCCACCTGCCTTGGCTTCCCAAACTGCTGGGAGTACA
GGCAAGCCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTCGCCCT
ATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCCGTGACTGGGAAAACCCTGGCGTTACCCAAC
TTAATCGCCTTGCAGCACATCCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCC
AACAGTTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTA
CGCGCAGCGTGACCGCTACACTTGCCAGCGCC

>BD43-14 (191, 100, 16q24.2) withM13R

Figure 3 Continued

GGAGGGTGTTTGCACAATCTCAGCTCACCACAACCTTTTCCTGCTGGGTTCAAGTGATTATCCTGCCTCAACCTCC
CGACTAGCTGGGATTACAGGCATGCACCACCATGCCTGGCTAATTTTGTATTTTTAGCAGAGACAGTGTTTCTCCA
TGTTGGTGAGGCTGGTCTCAAACTCCCGACCTCAGGTGATCCGCCTGCCTCAGCCTCCCAAACTGCTGGGAGTACA
GGCAAGCCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTCGCCCT
ATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACT
TAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAA
CAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAACGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGC
GCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGC

Figure 3 Continued